



TESIS - SS142501

PEMODELAN REGRESI KUANTIL TERSENSOR

(Studi Kasus Pengeluaran Rumah Tangga Untuk Konsumsi Rokok)

CINTIANI
NRP. 1315201006

DOSEN PEMBIMBING :
Dr. Dra. Ismaini Zain, M.Si.
Dr. Drs. Agus Suharsono, MS.

PROGRAM MAGISTER
JURUSAN STATISTIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2017



THESIS - SS142501

**CENSORED QUANTILE REGRESSION MODELLING
(A CASE STUDY OF HOUSEHOLD EXPENDITURE FOR
CONSUMPTION OF CIGARETTES)**

CINTIANI
NRP. 1315201006

SUPERVISOR :
Dr. Dra. Ismaini Zain, M.Si.
Dr. Drs. Agus Suharsono, MS.

PROGRAM OF MAGISTER
DEPARTMENT OF STATISTICS
FACULTY OF MATHEMATICS AND NATURAL SCIENCES
INSTITUT TEKNOLOGI SEPULUH NOPEMBER
SURABAYA
2017

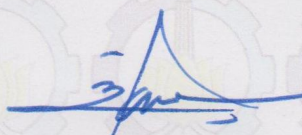
PEMODELAN REGRESI KUANTIL TERSENSOR (Studi Kasus Pengeluaran Rumah Tangga Untuk Konsumsi Rokok)

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar Magister Sains (M.Si)
di
Institut Teknologi Sepuluh Nopember
Oleh:

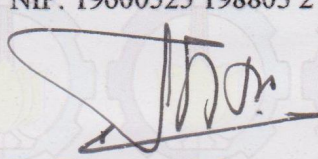
CINTIANI
NRP. 1315 201 006

Tanggal Ujian : 16 Januari 2017
Periode Wisuda : Maret 2017

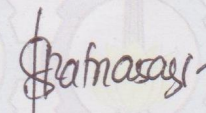
Disetujui oleh:


1. Dr. Dra. Ismaini Zain, M.Si.
NIP. 19600525 198803 2 001

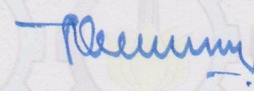
(Pembimbing I)


2. Dr. Drs. Agus Suharsono, MS
NIP. 19580823 198403 1 003

(Pembimbing II)


3. Dr. Vita Ratnasari, S.Si., M.Si.
NIP. 19700910 199702 2 001

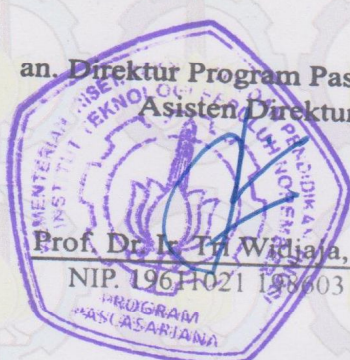
(Penguji)


4. Dr. Muhammad Mashuri, MT
NIP. 19620408 198701 1 001

(Penguji)

an. Direktur Program Pascasarjana
Asisten Direktur

Direktur Program Pasca Sarjana,


Prof. Dr. Ir. Tri Widjaja, M.Eng.
NIP. 19611021 198603 1 001

Prof. Ir. Djauhar Manfaat, M.Sc., Ph.D
NIP. 19601202 198701 1 001

PEMODELAN REGRESI KUANTIL TERSENSOR

(Studi Kasus Pengeluaran Rumah Tangga Untuk Konsumsi Rokok)

Nama Mahasiswa : Cintiani
NRP : 1315201006
Dosen Pembimbing : Dr. Dra. Ismaini Zain, M.Si
Dr. Drs. Agus Suharsono, MS

ABSTRAK

Data tersensor adalah data yang memuat nilai nol pada sebagian observasinya sedangkan sebagian nilai lainnya memiliki nilai tertentu yang bervariasi. Dalam analisis regresi seringkali terjadi pelanggaran normalitas pada saat data mengandung pencilan yang menyebabkan bentuk sebaran data tidak lagi simetrik. Akibatnya untuk metode kuadrat terkecil kurang tepat untuk melakukan analisis data yang tidak simetris, maka berkembanglah metode regresi median. Metode regresi median dilakukan dengan pendekatan *Least Absolute Deviation* (LAD) yang dikembangkan dengan mengganti rata-rata (*mean*) pada OLS menjadi median. Regresi kuantil tersensor digunakan pada kondisi terdapat data tersensor dan data yang memiliki sebaran data yang tidak simetrik. Estimator model regresi kuantil tersensor bisa diperoleh menggunakan solusi minimasi metode pemrograman linear dengan penggunaan tiga tahap algoritma. Tahapan dalam algoritma ini adalah dengan cara pemisahan dari probabilitas tersensor dan melakukan dua kali estimasi menggunakan regresi kuantil. Hasil dari estimasi pertama adalah mendapatkan sub sampel yang sesuai, sedangkan hasil dari estimasi yang kedua adalah mendapatkan estimasi yang efisien.

Dari hasil analisis yang digunakan pada data pengeluaran rumah tangga untuk konsumsi rokok diperoleh informasi bahwa pengeluaran rumah tangga untuk konsumsi rokok bervariasi antar kuantil, hal ini menunjukkan bahwa penggunaan model regresi kuantil tersensor sudah tepat digunakan dalam pemodelan data tersebut. Variabel yang memiliki pengaruh yang besar pada tingginya pengeluaran konsumsi rokok baik di wilayah gabungan (perkotaan dan pedesaan) adalah pendapatan, tingkat pendidikan kepala rumah tangga, jumlah anggota rumah tangga, rata-rata pengeluaran per kapita, dan jenis kelamin kepala rumah tangga. Sedangkan variabel yang memberikan efek yang semakin kecil seiring kenaikan pengeluaran konsumsi rokok adalah umur, sektor pekerjaan, dan wilayah tempat tinggal. Selain itu dari hasil simulasi diketahui bahwa regresi kuantil tersensor memiliki nilai RMSE yang cenderung lebih kecil dibandingkan dengan regresi kuantil.

Kata kunci: Konsumsi rokok, Least Absolute Deviation, Regresi kuantil tersensor, Tiga tahap algoritma

(halaman ini sengaja dikosongkan)

CENSORED QUANTILE REGRESSION MODELLING (A CASE STUDY OF HOUSEHOLD EXPENDITURE FOR CONSUMPTION OF CIGARETTES)

Name : Cintiani
NRP : 1315201006
Supervisor : Dr. Dra. Ismaini Zain, M.Si
Co Supervisor : Dr. Drs. Agus Suharsono, MS

ABSTRACT

Censored data is data that contains zero values in some observations while some other value has a specific value that varies. In regression analysis is often a violation of normality when the data contains outliers that caused a form of data distribution is no longer symmetric. As a result of the least squares method is less appropriate to perform data analysis that is not symmetrical, then developed median regression method. Median regression methods to do with the pproach Least Absolute Deviation (LAD) that developed by replacing the average (mean) in the OLS into the median. Censored quantile regression is used on the condition the data are censored and have symmetric distribution. Estimator censored quantile regression model can be obtained minimization solution using linear programming methods with the use of a three-stage algorithm. Stages in the algorithm is by means of separation of probability censored and did two times estimated using quantile regression. Results from the first estimate is getting a sub-sample is appropriate, while the results from the second estimate is obtained estimates that efficient.

From the results of data analysis used in household expenditure on cigarettes consumption was obtained that household expenditure on cigarettes consumption varies between quintile, this suggests that the use of censored quantile regression is appropriate used in model. Variables that had a great influence on the high cigarette consumption expenditure both in the urban and rural are income, education level of the head of household, number of household members, average expenditure per capita, and gender of household head. While the variables that give effect gets smaller as the increase in cigarette consumption expenditures are age, employment sector, and region of residence. In addition, from the simulation result that the censored quantile regression has RMSE values that tend to be smaller than the quantile regression.

Keywords: Censored Quantil Regression, Cigarettes Expenditure, Least Absolute Deviation, , Three-stage algorithm

(halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji dan syukur penulis hadirkan kepada Allah SWT, karena atas segala rahmat dan ridho-Nya sehingga tesis yang diberi judul **“Pemodelan Regresi Kuantil Tersensor (Studi Kasus Pengeluaran Rumah Tangga untuk Konsumsi Rokok)”** ini bisa terselesaikan. Tesis ini disusun untuk memenuhi salah satu syarat untuk menyelesaikan pendidikan di Program Magister Sains (M.Si) di Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember.

Terselesainya tesis ini tidak terlepas dari dukungan berbagai pihak yang telah memberikan kontribusi kepada penulis. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan terika kasih sedalam-dalamnya kepada yang terhormat:

1. Suamiku tercinta Lettu Laut (T) Beni Kusnan Kasogi, S.St.Han yang kini telah berada di syurgaNya, terima kasih atas segala cinta dan kasih sayang yang tiada henti penulis dapatkan. Ibunda tercinta Hj. Tureci, Ayahanda H. Muklas, dan seluruh keluarga besar yang senantiasa mendoakan, menyemangati dan memotivasi penulis.
2. Ibu Dr. Dra. Ismaini Zain, M.Si dan Bapak Dr. Drs. Agus Suharsono, MS selaku dosen pembimbing, yang telah bersedia meluangkan waktu untuk memberikan bimbingan, saran, dan ilmu yang sangat bermanfaat dalam penyelesaian tesis ini.
3. Ibu Dr. Vita Ratnasari, S.Si., M.Si dan Bapak Dr.Muhammad Mashuri, MT selaku dosen penguji yang telah memberikan banyak saran dan masukan agar tesis ini menjadi lebih baik.
4. Bapak Dr. Suhartono, M.Sc. selaku Ketua Jurusan Statistika ITS dan Bapak Dr.rer.pol. Heri Kuswanto, M.Si. selaku Kaprodi Pascasarjana Statistika FMIPA ITS yang telah memberikan fasilitas dan kemudahan selama proses perkuliahan.
5. Bapak Prof. Dr. I Nyoman Budiantara, M.Si., selaku dosen wali selama penulis menempuh perkuliahan.

6. Bapak /Ibu dosen pengajar di Jurusan Statistika FMIPA ITS yang telah memberikan ilmu yang bermanfaat.
7. Bapak/Ibu staf dan karyawan di Jurusan Statistika FMIPA ITS yang telah memberikan pelayanan dan fasilitas selama perkuliahan.
8. Tutus, Ifa, Rizfani, Asmita, Rani, Maman, Surya, serta rekan-rekan seperjuangan Magister Statistika angkatan 2015 lainnya, terima kasih atas saran, kerjasama dan kebersamaannya.
9. Serta, semua pihak yang telah membantu penulis, namun tidak dapat penulis sebutkan satu per satu.

Penulis menyadari bahwa tesis ini masih jauh dari sempurna, sehingga kritik dan saran sangat diharapkan. Semoga tesis ini dapat memberikan manfaat guna memperluas wawasan keilmuan pembacanya. Aamiin.

Surabaya, Januari 2017

Penulis

DAFTAR ISI

| | Halaman |
|---|-------------|
| HALAMAN JUDUL | i |
| HALAMAN PENGESAHAN | v |
| ABSTRAK..... | vii |
| ABSTRACT | ix |
| KATA PENGANTAR..... | xi |
| DAFTAR ISI..... | xiii |
| DAFTAR TABEL | xv |
| DAFTAR GAMBAR..... | xvii |
| DAFTAR LAMPIRAN | xix |
| | |
| BAB 1 PENDAHULUAN..... | 1 |
| 1.1 Latar Belakang | 1 |
| 1.2 Rumusan Masalah | 6 |
| 1.3 Tujuan Penelitian..... | 7 |
| 1.4 Manfaat Penelitian..... | 7 |
| 1.4 Batasan Masalah..... | 7 |
| | |
| BAB 2 TINJAUAN PUSTAKA | 9 |
| 2.1 Regresi Klasik | 9 |
| 2.2 Regresi Tersensor | 10 |
| 2.3 Regresi Kuantil..... | 13 |
| 2.4 Regresi Kuantil Tersensor..... | 17 |
| 2.6 Root Mean Square Error (RMSE) | 20 |
| 2.7 Tinjauan Non Statistika | 20 |
| | |
| BAB 3 METODE PENELITIAN..... | 23 |
| 3.1 Sumber Data..... | 23 |
| 3.2 Variabel Penelitian | 22 |
| 3.3 Metode Penelitian | 25 |

| | |
|---|-----------|
| 3.3.1 Estimasi Regresi Kuantil Tersensor..... | 25 |
| 3.3.2 Perbandingan Metode Regresi Kuantil Tersensor..... | 26 |
| 3.3.3 Penerapan Model Regresi Kuantil Tersensor..... | 27 |
| BAB 4 HASIL DAN PEMBAHASAN..... | 29 |
| 4.1 Estimasi Regresi Kuantil Tersensor | 29 |
| 4.2 Perbandingan Metode Regresi Kuantil Tersensor..... | 32 |
| 4.3 Penerapan Model Regresi Kuantil Tersensor..... | 34 |
| 4.3.1 Deskriptif Data Penelitian..... | 36 |
| 4.3.2 Hasil Estimasi Parameter Regresi Kuantil Tersensor..... | 41 |
| 4.3.3 Model Regresi Kuantil Tersensor..... | 49 |
| BAB 5 KESIMPULAN DAN SARAN..... | 55 |
| 5.1 Kesimpulan | 55 |
| 5.2 Saran | 57 |
| DAFTAR PUSTAKA..... | 59 |
| LAMPIRAN-LAMPIRAN..... | 63 |
| BIOGRAFI PENULIS | 79 |

DAFTAR TABEL

| | Halaman |
|--|---------|
| Tabel 2.1 Penelitian Sebelumnya | 22 |
| Tabel 3.1 Struktur Data Untuk Analisis | 25 |
| Tabel 4.1 RMSE Intersep dari Estimator Regresi Kuantil Tersensor dan Regresi Kuantil..... | 33 |
| Tabel 4.2 Statistik Deskriptif Pengeluaran Rumah Tangga untuk Konsumsi Rokok..... | 37 |
| Tabel 4.3 Statistik Deskriptif Variabel Prediktor yang Bersifat Kontinu | 37 |
| Tabel 4.4 Estimasi Parameter Regresi Kuantil Tersensor di Wilayah Gabungan | 42 |
| Tabel 4.5 Estimasi Parameter Regresi Kuantil Tersensor di Wilayah Perkotaan..... | 43 |
| Tabel 4.6 Estimasi Parameter Regresi Kuantil Tersensor di Wilayah Pedesaan..... | 44 |
| Tabel 4.7 Estimasi Parameter Regresi Kuantil di Wilayah Gabungan | 45 |
| Tabel 4.8 Estimasi Parameter Regresi Kuantil di Wilayah Perkotaan..... | 46 |
| Tabel 4.9 Estimasi Parameter Regresi Kuantil di Wilayah Pedesaan..... | 47 |
| Tabel 4.10 RMSE Intersep dari Estimator Regresi Kuantil Tersensor dan Regresi Kuantil | 48 |

(halaman ini sengaja dikosongkan)

DAFTAR GAMBAR

| | Halaman |
|---|---------|
| Gambar 2.1 y_i variabel normal dan y_i^* variabel tersensor..... | 11 |
| Gambar 2.2 Regresi Kuantil Fungsi ρ | 16 |
| Gambar 4.1 RMSE Intersep dari Estimator Regresi Kuantil Tersensor dan Regresi Kuantil..... | 33 |
| Gambar 4.2 <i>Scatter Plot</i> | 34 |
| Gambar 4.3 Diagram Pie Kategori dalam Variabel Respon | 36 |
| Gambar 4.4 Diagram Pie Kategori dalam Variabel Respon di Wilayah Perkotaan dan Pedesaan..... | 36 |
| Gambar 4.5 Diagram Pie Wilayah Tempat Tinggal | 39 |
| Gambar 4.6 Diagram Pie Jenis Kelamin di Wilayah Perkotaan dan Pedesaan | 39 |
| Gambar 4.7 Diagram Pie Pendidikan di Wilayah Perkotaan dan Pedesaan | 40 |
| Gambar 4.8 Diagram Pie Sektor Pekerjaan di Wilayah Perkotaan dan Pedesaan | 40 |

(halaman ini sengaja dikosongkan)

DAFTAR LAMPIRAN

| | Halaman |
|---|---------|
| Lampiran 1 Data Penelitian..... | 63 |
| Lampiran 2 Penjelasan Mengenai <i>Loss Function</i> | 64 |
| Lampiran 3 Syntax <i>Scatter Plot</i> | 66 |
| Lampiran 4 Syntax Perbandingan Metode..... | 68 |
| Lampiran 5 Command Stata untuk Regresi Kuantil Tersensor..... | 70 |
| Lampiran 6 Output Analisis Regresi Kuantil Tersensor..... | 71 |

(halaman ini sengaja dikosongkan)

BAB 1

PENDAHULUAN

1.1 Latar Belakang

Data tersensor merupakan data yang memuat nilai nol pada sebagian observasinya sedangkan untuk sebagian nilai lainnya memiliki nilai tertentu yang bervariasi. Ciri lain dari data tersensor adalah sebagian nilai dari suatu rentang tertentu ditransformasikan sebagai suatu nilai tunggal atau konstanta (Greene, 2008). Bila variabel tak bebas menunjukkan skala campuran, yaitu sebagian diskrit dan sebagian lagi kontinyu maka data yang demikian bisa dikatakan data yang tersensor (Zain dan Suhartono, 1997; Greene, 2008). Seringkali ditemui dalam suatu pemodelan regresi memiliki variabel respon yang mengandung nilai nol atau melewati batasan (*limit*) tertentu yang bukan nol, maka data ini dianggap sebagai outlier dan dibuang dalam pemodelan. Disisi lain akan sangat merugikan untuk membuang nilai-nilai variabel respon yang tidak masuk dalam *limit*, ketika informasi tersebut tersedia.

Model regresi tersensor adalah pendekatan untuk mengatasi data tersensor (Greene, 2008). Penggunaan regresi tersensor akan mengurangi efek bias karena data yang bernilai konstan dapat diolah secara bersamaan dengan data kontinu sehingga tidak akan kehilangan informasi yang berasal dari data diskrit. Dengan kata lain model regresi tersensor dapat mengakomodasi semua observasi, baik yang bernilai nol maupun tidak nol. Model regresi tersensor telah banyak dikembangkan dan digunakan dalam berbagai bidang penelitian diantaranya adalah Suhardi dan Llewelyn (2001) yang menggunakan regresi tersensor untuk menganalisa faktor-faktor yang berpengaruh terhadap kepuasan konsumen untuk jasa pengangkutan barang. Faidah (2012) memodelkan data tingkat Pengangguran Terbuka (TPT) perempuan di pulau Jawa menggunakan model tersensor dengan aspek wilayah. Permana (2013) memanfaatkan regresi tersensor dalam penelitian pengeluaran konsumsi rokok kota Kediri tahun 2011. Selain itu Cahyaningsih (2011) yang membandingkan model tersensor dan double hurdle dalam pemodelan pengeluaran konsumsi rokok di Kalimantan Timur.

Estimasi parameter regresi tersensor menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan memaksimumkan fungsi likelihood sehingga diperoleh penaksir yang konsisten dan efisien untuk sampel yang berukuran besar. Metode MLE berbasis *conditional mean* yaitu estimator yang diperoleh menitikberatkan pada *mean* dari distribusi variabel respon. Nilai *mean* menunjukkan ukuran pemusatan dari suatu distribusi sehingga hanya sedikit informasi yang diketahui dari keseluruhan distribusi. Oleh sebab itu pendekatan dengan metode ini hanya mampu menduga model dari fungsi bersyarat *mean* dan tidak mewakili keseluruhan data dari distribusi (Hao dan Naiman, 2007). Terdapat metode lain yang mampu menggambarkan hubungan antara variabel prediktor terhadap variabel respon pada berbagai titik kuantil (*conditional quantile*). Metode ini dapat memberikan hasil yang tepat dan stabil pada kehadiran pencilan serta dapat membatasi pengaruh dari pencilan (Furno, 2014).

Analisis model regresi kuantil pertama kali diperkenalkan oleh Koenker dan Basset (1978) dan dapat digunakan pada kondisi data yang heterogen. Untuk mendapatkan estimator parameter model regresi kuantil didapatkan dengan metode pemrograman linier diantaranya menggunakan algoritma simpleks, *interior-point*, dan *smoothing*. Menurut Chen dan Wei (2005), ketiga algoritma tersebut memiliki kekurangan dan kelebihan masing-masing. Algoritma simpleks memberikan hasil yang lambat pada jumlah data observasi yang besar ($n > 100.000$) namun merupakan algoritma yang paling stabil dibandingkan dengan algoritma *interior-point* dan *smoothing*. Algoritma simpleks dapat memberikan solusi pada beberapa jenis data terutama pada data dengan jumlah outlier yang besar. Algoritma *interior-point* memberikan hasil yang sangat cepat pada data yang ramping, dimana memiliki jumlah observasi yang besar ($n > 100.000$) dan jumlah kovariat yang kecil. Algoritma ini memiliki struktur yang sederhana dan dapat diadaptasi pada berbagai situasi seperti regresi kuantil. Sedangkan algoritma *smoothing* memiliki teori yang sederhana untuk regresi kuantil dan memiliki kelebihan dalam kecepatan komputasi pada jumlah kovariat yang besar.

Model regresi kuantil tersensor diperkenalkan oleh Powell (1986), sehingga dikenal sebagai estimator Powell. Estimator regresi kuantil tersensor

bisa diperoleh menggunakan solusi minimasi metode pemrograman linear. Buchinsky (1994) menawarkan penggunaan *Iterative Linear Programming Algorithm* (ILPA) yang melibatkan *Barrodale-Robert Algorithm* (BRA). Namun, metode ILPA memiliki kekurangan bahwa tidak ada kepastian konvergensi tercapai dan sekalipun tercapai, hal ini tidak menjamin solusi yang dihasilkan merupakan *local minima* dari permasalahan optimasi regresi kuantil tersensor. Selanjutnya Fitzenberger (1997) mengembangkan algoritma BRCENS sebagai adaptasi dari algoritma BRA dalam penjaminan konvergensi *local optima* yang dibangun atas karakteristik estimasi menggunakan *Interpolation Property*.

Simulasi studi yang dilakukan oleh Fitzenberger (1997) memperlihatkan bahwa algoritma BRCENS memberikan hasil yang lebih baik dari algoritma ILPA. Namun demikian seluruh algoritma tersebut memberikan performa yang kurang baik pada kondisi proporsi data tersensor yang besar. Dalam mengatasi permasalahan tersebut, dikembangkan tiga tahap algoritma yang dikenalkan oleh Chernozhukov dan Hong (2002) dimana tahapan yang digunakan lebih sederhana, robust, dan lebih dekat dengan titik sensoring.

Beberapa penelitian yang menggunakan model regresi kuantil tersensor diantaranya adalah Chernozhukov dan Hong (2002) yang menggunakan pendekatan tiga step regresi kuantil tersensor dalam kasus perselingkuhan di Amerika Serikat, sedangkan untuk mengevaluasi kebaikan performa model atau estimator digunakan RMSE. Gustavsen *et al.* (2008) menggunakan metode regresi kuantil tersensor dalam penelitian mengenai efek jumlah pembelian es krim dalam meningkatkan Pajak Pertambahan Nilai (PPN) untuk makanan kurang sehat dan menghilangkan PPN untuk makanan sehat dan menggunakan pseudo R^2 untuk mengevaluasi kebaikan performa modelnya. Selanjutnya Gustavsen dan Rickertsen (2013) melakukan pendekatan regresi kuantil tersensor dalam penyesuaian tarif PPN untuk mempromosikan diet sehat di Norwegia. Lusiana (2015) melakukan penelitian mengenai model kuantil tersensor bayesian pada kasus pengeluaran rumah tangga untuk konsumsi susu dan menggunakan RMSE untuk mengevaluasi kebaikan performa modelnya.

Indonesia merupakan salah satu negara dengan konsumsi tembakau terbesar di dunia. Berdasarkan survey World Health Organization (WHO) pada tahun 2011, Indonesia merupakan negara dengan jumlah perokok aktif ketiga terbesar di dunia setelah Cina dan India. Statistik konsumsi rokok masyarakat Indonesia tersebut sejalan dengan tingginya prevalensi merokok di tanah air. Hasil *Global Adult Tobacco Survey* (GATS) pada 2011 memperlihatkan bahwa jumlah pengguna tembakau, baik berupa rokok maupun penggunaan lainnya tanpa asap (*smokeless form*), mencapai 61 juta orang atau mencakup sekitar 36 persen dari total penduduk Indonesia.

Rokok mengandung lebih dari 4000 jenis bahan kimia berbahaya bagi kesehatan, mulai dari nikotin maupun zat lainnya yang bisa menyebabkan kanker dan zat beracun bagi tubuh. Menurut Dinas Kesehatan Kabupaten Kudus (2015) bahaya merokok bagi kesehatan bukan saja bagi perokok tetapi bagi orang sekitar karena efek asap rokok atau perokok pasif. Berikut beberapa penyakit berbahaya yang diakibatkan oleh rokok bagi kesehatan tubuh:

1. Penyakit paru-paru

Efek dari perokok yang paling pertama merusak organ tubuh akibat asap rokok adalah paru-paru. Asap rokok tersebut terhirup dan masuk ke dalam paru-paru sehingga menyebabkan paru-paru mengalami radang, bronchitis, pneumonia. Bahaya dari zat nikotin menyebabkan kerusakan sel-sel dalam organ paru-paru yang bisa berakibat fatal yaitu kanker paru-paru. Bahaya merokok bagi kesehatan ini tentu sangat beresiko dan bisa menyebabkan kematian.

2. Penyakit impotensi dan organ reproduksi

Efek bahaya merokok bagi kesehatan lainnya adalah bisa mengakibatkan impotensi, kasus seperti ini sudah banyak dialami oleh para perokok. Sebab kandungan bahan kimia yang sifatnya beracun tersebut bisa mengurangi produksi sperma pada pria. Bukan hanya itu saja, pada pria juga bisa terjadi kanker di bagian testis. Efek bahaya merokok bagi kesehatan remaja bisa menyebabkan resiko tidak memiliki keturunan. Sedangkan pada wanita yang merokok, efek dari rokok bisa mengurangi tingkat kesuburan wanita.

3. Penyakit lambung

Hal yang terlihat sepele ketika menghisap rokok adalah aktifitas otot di bawah kerongkongan semakin meningkat. Otot sekitar saluran pernafasan bagian bawah akan lemah secara perlahan sehingga proses pencernaan menjadi terhambat. Bahaya merokok bagi kesehatan juga bisa dirasakan sampai ke lambung, karena asap rokok yang masuk ke sistem pencernaan akan menyebabkan meningkatnya asam lambung. Jika hal ini dibiarkan terus menerus maka bukan tidak mungkin akan menjadi penyakit yang lebih kronis seperti tukak lambung yang lebih sulit diobati.

4. Resiko stroke

Pada perokok aktif bisa saja menderita serangan stroke, karena efek samping rokok bisa menyebabkan melemahnya pembuluh darah. Ketika pelemahan tersebut terjadi dan kerja pembuluh darah terhambat bisa menyebabkan serangan radang di otak. Hal itulah yang bisa beresiko terjadi stroke meskipun orang tersebut tidak ada latar belakang darah tinggi atau penyakit penyebab stroke lainnya. Penyebab stroke tersebut bersumber dari kandungan kimia berbahaya seperti nikotin, tar, karbon monoksida dan gas oksidan yang terkandung dalam rokok.

Dengan melihat banyaknya bahaya yang ditimbulkan akibat konsumsi rokok, menyebabkan biaya yang harus dibayar akibat dampak buruk tersebut sangat mahal. Banyak masyarakat Indonesia yang bukan perokok setiap hari terpapar asap rokok sehingga berisiko menderita berbagai penyakit yang ditimbulkan oleh asap rokok. Pada saat yang sama, biaya kesehatan yang dikeluarkan untuk berbagai penyakit yang dikaitkan dengan penggunaan tembakau menjadi sangat tinggi setiap tahun. Fakta ini sejatinya memberi konfirmasi bahwa kerugian yang ditimbulkan oleh rokok lebih besar daripada manfaat ekonomi yang dihasilkan.

Persoalan semakin rumit karena prevalensi merokok pada masyarakat miskin ternyata juga sangat tinggi. Hal itu tercermin dari tingginya pengeluaran penduduk miskin yang dialokasikan untuk membeli rokok. Hasil perhitungan Badan Pusat Statistik (BPS) memperlihatkan, sumbangan pengeluaran untuk

rokok terhadap garis kemiskinan menempati posisi kedua setelah pengeluaran untuk beras. BPS mencatat, pada Maret 2015, kontribusi pengeluaran untuk rokok terhadap garis kemiskinan mencapai 8,24 persen di perkotaan dan 7,07 persen di pedesaan, jauh lebih tinggi dibanding kontribusi pengeluaran untuk pendidikan yang hanya sebesar 2,46 persen di perkotaan dan 1,39 persen di pedesaan. Artinya masyarakat miskin lebih banyak menghabiskan uang untuk rokok daripada kebutuhan pendidikan.

Dengan fakta tingginya angka konsumsi rokok serta bahaya yang ditimbulkan dari konsumsi rokok, maka peneliti akan mencoba melihat faktor-faktor yang mempengaruhi pengeluaran rumah tangga untuk konsumsi rokok di Indonesia. Studi awal menunjukkan bahwa faktor-faktor yang mempengaruhi pengeluaran rumah tangga untuk konsumsi rokok di Indonesia memiliki data yang menyebar. Hal ini bisa terlihat dari diagram pencar antara variabel tak bebas dengan masing-masing variabel bebas. Selain itu, variabel tak bebas menunjukkan skala campuran, yaitu sebagian diskrit dan sebagian lagi kontinyu sehingga dapat dikatakan sebagai data tersensor. Terkait dengan hal tersebut, maka regresi kuantil tersensor cocok diterapkan untuk menganalisis faktor-faktor yang mempengaruhi pengeluaran rumah tangga untuk konsumsi rokok di Indonesia.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dipaparkan, dapat dirumuskan beberapa masalah berikut.

1. Bagaimana bentuk estimasi parameter model regresi kuantil tersensor?
2. Bagaimana perbandingan performa metode regresi kuantil tersensor dibandingkan dengan regresi kuantil ?
3. Bagaimana mengaplikasikan model regresi kuantil tersensor pada faktor-faktor yang mempengaruhi pengeluaran rumah tangga untuk konsumsi rokok?

1.3 Tujuan Penelitian

Berikut ini merupakan tujuan yang ingin dicapai dalam penelitian ini.

1. Mendapatkan bentuk estimasi parameter model regresi kuantil tersensor.
2. Melakukan perbandingan performa metode regresi kuantil tersensor dengan regresi kuantil.
3. Memodelkan regresi kuantil tersensor pada faktor-faktor yang mempengaruhi pengeluaran rumah tangga untuk konsumsi rokok.

1.4 Manfaat Penelitian

Manfaat dari hasil penelitian ini nantinya sebagai berikut.

1. Mengembangkan wawasan dan pengetahuan mengenai analisis regresi kuantil pada umumnya dan regresi kuantil tersensor pada khususnya.
2. Memberikan informasi yang lebih lengkap tentang model pengeluaran rumah tangga untuk konsumsi rokok, sehingga diharapkan bisa membantu upaya pengambil kebijakan di suatu wilayah.

1.5 Batasan Masalah

Beberapa batasan masalah dalam penelitian ini adalah sebagai berikut.

1. Ukuran untuk mengevaluasi performa estimator adalah RMSE.
2. Pada saat memodelkan regresi kuantil, data yang digunakan adalah data asli (tanpa mengeluarkan data tersensor).
3. Kuantil yang akan dimodelkan pada model gabungan (wilayah pedesaan dan perkotaan) adalah 0,1, 0,25, 0,5, 0,75, dan 0,9, sedangkan pada model di wilayah pedesaan dan perkotaan adalah 0,25, 0,5, 0,75, dan 0,9.

(halaman ini sengaja dikosongkan)

BAB 2

TINJAUAN PUSTAKA

2.1 Regresi Klasik

Regresi merupakan suatu metode untuk mengetahui hubungan antara satu variabel respon terhadap variabel lainnya (prediktor). Salah satu metode yang bisa digunakan untuk mendapatkan estimasi regresi adalah *Ordinary least square* (OLS) atau sering disebut sebagai metode regresi klasik. Metode OLS mendefinisikan estimasi parameter sebagai suatu nilai yang meminimumkan jumlah kuadrat antara pengamatan dan model. Persamaan umum regresi dengan k -variabel dinyatakan dalam persamaan (2.1) berikut ini (Gujarati, 2004).

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_k x_{ki} + \varepsilon_i, i = 1, 2, \dots, n \quad (2.1)$$

Penjabaran dalam bentuk matriks, persamaan (2.1) dapat ditulis sebagai persamaan (2.2).

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{1,1} & x_{2,1} & \cdots & x_{k,1} \\ 1 & x_{1,2} & x_{2,2} & \cdots & x_{k,2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1,n} & x_{2,n} & \cdots & x_{k,n} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (2.2)$$

Secara umum, persamaan (2.2) dapat dituliskan dalam model regresi linear (k -variabel) menjadi persamaan (2.3).

$$y_i = \mathbf{x}_i^T \boldsymbol{\beta} + \varepsilon_i \quad (2.3)$$

dengan :

$$y_i = (y_1 \ y_2 \ \cdots \ y_n)^T \text{ dengan vektor ukuran } n \times 1$$

$$\mathbf{x}_i^T = (1 \ x_{1i} \ x_{2i} \ \cdots \ x_{ki}) \text{ dengan matriks ukuran } n \times (k + 1)$$

$$\boldsymbol{\beta} = (\beta_0 \ \beta_1 \ \beta_2 \ \cdots \ \beta_k)^T \text{ dengan matriks ukuran } (k + 1) \times 1$$

$$\boldsymbol{\varepsilon} = (\varepsilon_1 \ \varepsilon_2 \ \cdots \ \varepsilon_n)^T \text{ dengan vektor ukuran } n \times 1$$

di mana:

$$y_i = \text{Variabel respon dengan pengamatan ke-} i \text{ yang bersifat random}$$

- \mathbf{x}_i^T = Variabel prediktor ke- j dan pengamatan ke- i
 β = Parameter regresi pada variabel ke- j
 ε = *Error* random model regresi pada pengamatan ke- i

2.2 Regresi Tersensor

Model regresi tersensor atau model regresi tobit pertama kali diperkenalkan oleh James Tobin pada tahun 1958 yang digunakan untuk menganalisis pengeluaran rumah tangga untuk membeli mobil di Amerika Serikat. Pada awalnya model regresi tersensor digunakan untuk menganalisis pengeluaran rumah tangga terhadap barang mewah atau barang tahan lama lainnya yang seringkali bernilai nol atau melewati nilai batasan (*limit*) tertentu yang bukan nol. Dalam kondisi variabel respon semacam ini maka keberadaan nilai atau observasi *limit* tersebut harus mendapatkan perhatian ketika dilakukan estimasi parameter regresi diantara variabel respon dengan variabel-variabel prediktor. Serta terhadap pengujian hipotesis atas hubungan kedua variabel tersebut. Apabila nilai respon diluar *limit* diabaikan maka analisis probit dapat digunakan dalam model. Disisi lain akan sangat merugikan untuk membuang nilai-nilai variabel respon yang tidak masuk dalam *limit*, ketika informasi tersebut tersedia. Oleh karena itu, model regresi tersensor diharapkan dapat mengatasi permasalahan tersebut (Tobin, 1958).

Pemodelan regresi tersensor diawali dengan memperhatikan model berikut (Greene, 2008) :

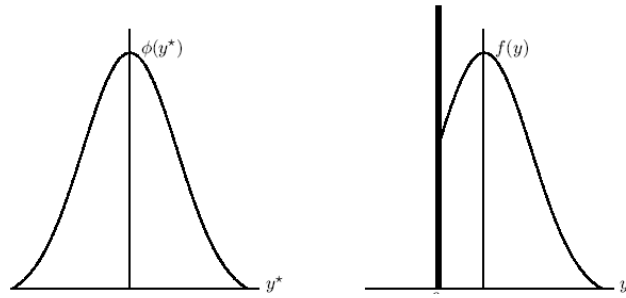
$$y_i = \mathbf{x}_i^T \beta + \varepsilon_i \quad (2.4)$$

dengan y_i adalah variabel dependen yang diobservasi untuk nilai yang lebih besar dari τ dan tersensor untuk nilai lainnya. \mathbf{x}_i^T adalah vektor variabel bebas $\mathbf{x}_i^T = [1 \quad X_{1i} \quad \cdots \quad X_{pi}]$, β adalah vektor parameter koefisien $\beta = [\beta_0 \quad \beta_1 \quad \cdots \quad \beta_p]^T$, dan ε_i adalah *error* yang diasumsikan berdistribusi $N(0, \sigma^2)$.

Misalkan terdapat sebanyak n data observasi yang terdiri atas satu variabel respon dan p variabel prediktor maka variabel respon dikatakan tersensor pada batas bawah apabila untuk setiap $i = 1, 2, \dots, n$ berlaku persamaan berikut (McBee, 2010),

$$y_i = \begin{cases} \tau & y_i^* \leq \tau \\ y_i^* & y_i^* > \tau \end{cases} \quad (2.5)$$

Dengan τ adalah suatu konstanta tertentu. Hal ini dapat diilustrasikan pada Gambar 2.1.



Gambar 2.1 y_i variabel normal dan y_i^* variabel tersensor

Leiker (2012) menyatakan bahwa persamaan (2.6) dapat ditulis menjadi persamaan (2.6).

$$y_i = \max(y_i^*, \tau) \quad (2.6)$$

Nilai observasi y_i^* dapat dinyatakan sebagai persamaan (2.7).

$$y_i^* = f(\mathbf{x}_i) + \varepsilon_i \quad (2.7)$$

di mana : $f(\mathbf{x}_i) = \mathbf{x}_i^T \boldsymbol{\beta}$

Dengan demikian diperoleh model regresi tersensor dalam persamaan (2.8) atau (2.9).

$$y_i = \begin{cases} \tau & y_i^* \leq \tau \\ \mathbf{x}_i^T \boldsymbol{\beta} + \varepsilon_i & y_i^* > \tau \end{cases} \quad (2.8)$$

$$y_i = \max(\mathbf{x}_i^T \boldsymbol{\beta} + \varepsilon_i, \tau) \quad (2.9)$$

di mana:

y_i = nilai variabel respon tersensor ke- i

y_i^* = nilai variabel respon laten ke- i

τ = nilai titik sensor yang diketahui

\mathbf{x}_i = $(1, x_{1i}, x_{2i}, \dots, x_{ki})$

$\boldsymbol{\beta}$ = $[\beta_0 \ \beta_1 \ \dots \ \beta_k]^T$ merupakan parameter regresi tersensor

ε_i = *error* model yang berdistribusi $N(0, \sigma^2)$

i = 1, 2, ..., n

Apabila model regresi tobit dengan variabel respon y tersensor pada batas atas, maka tanda pertidaksamaan pada persamaan (2.8) dan (2.9) dibuuh sebaliknya.

Diketahui bahwa y_i berdistribusi normal dengan rata-rata $\mathbf{x}_i^T \boldsymbol{\beta}$ dan varians σ^2 maka diperoleh probabilitas data tersensor dan tidak tersensor sesuai dengan persamaan (2.10) dan (2.11). Nilai ϕ adalah *probability distribution function* (pdf) dari distribusi normal standar dan Φ adalah *cumulative distribution function* (cdf) dari distribusi normal standar.

$$\begin{aligned} \Pr(y_i = 0) &= \Pr(y_i^* \leq 0) = \Pr(\varepsilon_i \leq -\mathbf{x}_i^T \boldsymbol{\beta}) = \Pr\left(\frac{\varepsilon_i}{\sigma} \leq -\frac{\mathbf{x}_i^T \boldsymbol{\beta}}{\sigma}\right) \\ &= \Phi\left(-\frac{\mathbf{x}_i^T \boldsymbol{\beta}}{\sigma}\right) = 1 - \Phi\left(\frac{\mathbf{x}_i^T \boldsymbol{\beta}}{\sigma}\right) \end{aligned} \quad (2.10)$$

$$\begin{aligned} \Pr(y_i > 0) &= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(y_i - \mathbf{x}_i^T \boldsymbol{\beta})^2\right) = \sigma^{-1} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(y_i - \mathbf{x}_i^T \boldsymbol{\beta})^2\right) \\ \text{misal } z &= \frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{\sigma} \text{ maka } \phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(z)^2\right) \\ \text{sehingga } \phi\left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{\sigma}\right) &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{\sigma}\right)^2\right) \\ \Pr(y_i > 0) &= \sigma^{-1} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2}(y_i - \mathbf{x}_i^T \boldsymbol{\beta})^2\right) = \sigma^{-1} \phi\left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{\sigma}\right) \end{aligned} \quad (2.11)$$

Fungsi likelihood model regresi tersensor dapat dituliskan pada persamaan (2.12).

$$L(\boldsymbol{\beta}) = \prod_{i=1}^n \left[1(y_i > 0) \sigma^{-1} \phi\left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{\sigma}\right) \right] \left[1(y_i = 0) \left\{ 1 - \Phi\left(\frac{\mathbf{x}_i^T \boldsymbol{\beta}}{\sigma}\right) \right\} \right] \quad (2.12)$$

Sedangkan fungsi log likelihood dari persamaan (2.12) adalah sebagai berikut.

$$l(\boldsymbol{\beta}) = \log L(\boldsymbol{\beta}) = \sum_{i=1}^n 1(y_i > 0) \log \left\{ \frac{1}{\sigma} \varphi \left(\frac{y_i - \mathbf{x}_i^T \boldsymbol{\beta}}{\sigma} \right) \right\} + 1(y_i = 0) \log \left\{ 1 - \Phi \left(\frac{\mathbf{x}_i^T \boldsymbol{\beta}}{\sigma} \right) \right\} \quad (2.13)$$

Nilai maksimum fungsi likelihood dapat diperoleh dari turunan pertama fungsi likelihoodnya terhadap parameter $\boldsymbol{\beta}$ kemudian disamadengankan nol dengan mengasumsikan σ^2 diketahui. Namun karena penyelesaian persamaan ini bersifat nonlinear maka sukar dilakukan perhitungan secara analitis. Sehingga digunakan metode iteratif Newton-Raphson (Greene, 2008).

2.3 Regresi Kuantil

Dalam analisis regresi seringkali terjadi pelanggaran normalitas pada saat data mengandung pencilan yang menyebabkan bentuk sebaran data tidak lagi simetrik. Akibatnya metode kuadrat terkecil kurang tepat untuk melakukan analisis data yang tidak simetris, maka berkembanglah metode regresi median. Metode regresi median dilakukan dengan pendekatan *Least Absolute Deviation* (LAD) yang dikembangkan dengan mengganti rata-rata (*mean*) pada OLS menjadi median dengan mempertimbangkan apabila data berbentuk lonceng tidak simetris. Seiring dengan perkembangan waktu, pendekatan regresi median juga kurang tepat digunakan karena hanya melihat pada dua kelompok data saja sehingga berkembanglah metode regresi kuantil yang bisa digunakan pada lebih dari dua kelompok data. Pendekatan metode regresi kuantil dengan memisahkan atau membagi data menjadi dua atau lebih kelompok dimana dicurigai adanya perbedaan nilai estimasi pada kuantil-kuantil tersebut.

Regresi kuantil merupakan suatu metode analisis regresi yang berguna dalam mengestimasi parameter dan tidak mudah terpengaruh oleh kehadiran pencilan. Selain itu, metode ini dapat memberikan hasil yang tepat dan stabil pada kehadiran pencilan (Furno, 2014). Metode ini dapat digunakan untuk menggambarkan hubungan satu atau beberapa variabel prediktor terhadap satu variabel respon pada berbagai titik kuantil (*conditional quantile*) dari distribusi variabel respon tersebut, sehingga dapat digunakan pada kondisi data yang heterogen. Hal ini berbeda dengan analisis regresi linear yang hanya dapat

menggambarkan hubungan sebab akibat pada *mean (conditional mean)* variabel respon (Koenker dan Hallock, 2001).

Metode regresi kuantil diperkenalkan pertama kali oleh Roger Koenker dan Gilbert Basset pada tahun 1978. Regresi kuantil merupakan perluasan model regresi linier. Regresi kuantil dapat memberikan gambaran yang menyeluruh tentang bagaimana variabel independen berhubungan dengan distribusi bersyarat yang mendasari variabel dependen, terutama ketika distribusi bersyarat adalah heterogen dan tidak mengikuti distribusi normal standar. Regresi Kuantil sangat berguna jika data tidak homogen (varian y berubah seiring perubahan X) dan tidak simetris, terdapat ekor pada sebaran atau *truncated distribution* (Koenker, 2005).

Pada regresi linier $E(y|\mathbf{x}_i) = \mathbf{x}_i^T \boldsymbol{\beta}$ sementara dalam regresi kuantil $Q_Y(y|\mathbf{x}_i) = \mathbf{x}_i^T \boldsymbol{\beta}_\theta$ yang dapat diuraikan menjadi:

$$y_i = \beta_{\theta,0} + \sum_{k=1}^p \beta_{\theta k} x_{ik} + \varepsilon_{\theta i}, i = 1, 2, \dots, n \quad (2.14)$$

dimana $\beta_{\theta 0}, \beta_{\theta 1}, \dots, \beta_{\theta p}$ merupakan parameter yang tidak diketahui pada kuantil ke- θ dan $\varepsilon_{\theta i}$ merupakan residual dari model regresi pada sampel sebanyak n dan pada kuantil ke- θ . Dalam bentuk matriks, regresi kuantil dapat ditulis sebagai berikut:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \begin{bmatrix} \beta_{\theta 0} \\ \beta_{\theta 1} \\ \vdots \\ \beta_{\theta k} \end{bmatrix} + \begin{bmatrix} \varepsilon_{\theta 1} \\ \varepsilon_{\theta 2} \\ \vdots \\ \varepsilon_{\theta n} \end{bmatrix} \quad (2.15)$$

Bentuk umum model regresi kuantil linier terdapat pada persamaan (2.16) berikut ini (Buhai, 2005).

$$y_i = \mathbf{x}_i^T \boldsymbol{\beta}_\theta + \varepsilon_\theta, \quad 0 < \theta < 1 \quad (2.16)$$

di mana:

y_i = nilai variabel respon ke- i

$\mathbf{x}_i^T = (1, x_{1i}, x_{2i}, \dots, x_{ki})$

β_θ = parameter model regresi pada kuantil ke- θ

ε_θ = *error* model regresi kuantil ke- θ

i = 1,2,...,n.

Estimasi parameter model regresi kuantil diawali dengan menyatakan fungsi peluang kumulatif dari variabel random Y seperti persamaan (2.17), sehingga kuantil ke θ dari variabel ini dapat dituliskan sebagaimana persamaan (2.18) berikut ini (Chen, 2005).

$$F(y) = P(Y \leq y) \quad (2.17)$$

$$Q_\theta(Y) = \inf \{y : F(y) \geq \theta\} \quad (2.18)$$

Kemudian pada kuantil ke- θ dari $F(y)$ dapat diperoleh dengan meminimumkan fungsi tersebut terhadap Q yaitu :

$$\theta \int_{y>q} |y - q| dF_Y(y) + (1 - \theta) \int_{y<q} |y - q| dF_Y(y) \quad (2.19)$$

Jika Y merupakan fungsi X yang diketahui memiliki fungsi probabilitas $F_{Y|X}(y)$, maka kuantil ke- θ dari fungsi tersebut dapat dituliskan menjadi $Q_{Y_i}(\theta|x)$ yang merupakan fungsi dari X dan diselesaikan dengan persamaan berikut:

$$\min_q \theta \int_{i=1, y>q}^n |y - q| dF_Y(y) + (1 - \theta) \int_{i=1, y<q}^n |y - q| dF_Y(y) \quad (2.20)$$

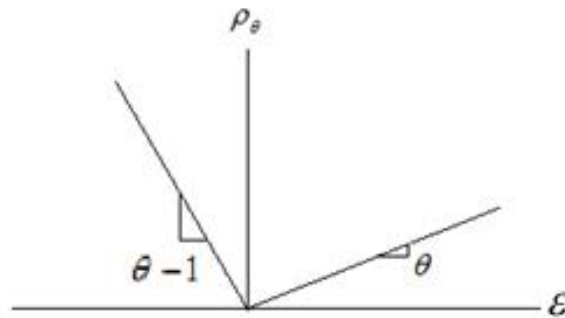
Dalam regresi kuantil terdapat fungsi kuantil bersyarat ke- θ yang mempertimbangkan penduga β_θ sehingga diperoleh solusi untuk permasalahan tersebut yang dinyatakan sebagai berikut:

$$\beta(\theta)^* = \min_\beta \left\{ \theta \sum_{i=1, y>x}^n |y - \mathbf{x}^T \boldsymbol{\beta}| + (1 - \theta) \sum_{i=1, y<x}^n |y - \mathbf{x}^T \boldsymbol{\beta}| \right\} \quad (2.21)$$

Salah satu metode estimasi parameter secara numerik untuk regresi kuantil yaitu menggunakan algoritma simpleks yang telah dikembangkan oleh Barrodale dan Robert pada tahun 1974. Algoritma simpleks memberikan hasil yang lambat pada jumlah data observasi yang besar ($n > 100.000$) namun merupakan algoritma yang paling stabil dibandingkan dengan algoritma *interior-point* dan *smoothing*. Algoritma simpleks dapat memberikan solusi pada beberapa jenis data terutama

pada data dengan jumlah outlier yang besar (Chen dan Wei, 2005). Algoritma simpleks merupakan cara untuk menentukan kombinasi optimal dari tiga variabel atau lebih. Algoritma tersebut memberikan solusi permasalahan program linear yang melibatkan banyak variabel keputusan dengan bantuan komputasi.

Regresi kuantil merupakan regresi alternatif yang dikembangkan oleh Koenker dan Basset (1978) karena regresi OLS hanya memberi masalah rata-rata. Estimasi parameter dalam regresi OLS, hanya dapat digunakan untuk memberi solusi permasalahan *mean*. Regresi kuantil merupakan pengembangan dari regresi kuantil median. Regresi OLS diestimasi dengan meminimumkan jumlah kuadrat residual, sedangkan regresi kuantil akan meminimumkan jumlah absolut residual yang lebih dikenal sebagai metode *Least Absolute Deviation* (Koenker, 1978).



Gambar 2.2 . Regresi Kuantil Fungsi ρ

Sebagai pengembangan dari regresi median, regresi kuantil meminimumkan $\sum_i^n |\varepsilon_i|$ dengan memberi pembobot yang berbeda. Pada regresi median, *error* diberikan bobot yang sama sementara pada regresi kuantil (selain kuantil ke 50%) *error* diberikan bobot θ untuk yang lebih dari sama dengan nol (*underprediction*) dan $(1 - \theta)$ untuk *error* yang kurang dari nol (*overprediction*), dengan θ adalah kuantil. Perkalian antara *error* dengan bobot tersebut kemudian disebut sebagai *loss function*.

Seperti pada metode OLS yang meminimumkan jumlah kuadrat residual untuk estimasi f , dengan metode *Least Absolute Deviation* (LAD) estimasi β dalam regresi kuantil pada persamaan (2.16) dilakukan dengan meminimumkan *loss function*.

Nilai β yang meminimumkan kuadrat error dengan metode OLS yaitu:

$$\hat{\beta}_{\theta} = \arg \min_{\beta_{\theta}} \left[\sum_{i=1}^n \rho_{\theta}(y_i - \mathbf{x}_i^T \beta) \right] \quad (2.22)$$

dimana $\rho_{\theta}(\varepsilon)$ merupakan *check function* dengan definisi $\rho_{\theta}(\varepsilon) = \varepsilon(\theta - I[\varepsilon < 0])$ atau:

$$\rho_{\theta}(\varepsilon) = \begin{cases} \theta \varepsilon & , \varepsilon \geq 0 \\ (\theta - 1) \varepsilon & , \varepsilon < 0 \end{cases}$$

dimana $\theta \in (0, 1)$, $\arg \min_{\beta_{\theta}}$ merupakan nilai yang meminimumkan nilai β dan $I(.)$ merupakan fungsi indikator.

Persamaan (2.20) tidak *close form*, maka metode iterasi numerik biasa tidak dapat digunakan untuk menyelesaikan persamaan tersebut. Sehingga digunakan metode pemrograman linier yaitu metode simpleks (Koenker dan Machado, 1999).

2.4 Regresi Kuantil Tersensor

Model regresi tersensor biasa digunakan untuk mengkoreksi data tersensor, dan sebagai patokan untuk menaksir efek dari rata-rata. Namun karena efek marginal mungkin berbeda pada kondisi kuantil yang lebih rendah ataupun lebih tinggi sebagai perbandingan dengan *conditional mean*, maka model regresi tersensor menjadi kurang akurat untuk digunakan. Untuk mengatasi kekurangan tersebut, lebih lanjut digunakan estimator regresi kuantil tersensor sebagai estimator konsisten dalam permasalahan distribusi *error* yang bersifat heteroskedastisitas dan tidak normal (Powell, 1986).

Diberikan y_i^* yang menunjukkan variabel respon yang bersifat laten diasumsikan dihasilkan dari model linier.

$$y_i^* = \mathbf{x}_i^T \beta_0 + \varepsilon_i, i = 1, 2, \dots, n \quad (2.23)$$

dimana ε_i berdistribusi independen dan identik (iid) dari fungsi distribusi F dengan fungsi densitas f. Variabel y_i^* tidak dapat diobservasi secara langsung karena merupakan variabel tersensor sehingga sebaliknya dilihat model

$y_i = \max\{0, y_i^*\}$. Analisis hubungan pada data tersensor terjadi jika variabel eksplanatori telah tersedia, tetapi nilai variabel dependen hanya diketahui pada observasi dimana variabel dependen lebih besar (tersensor kanan) atau lebih kecil (tersensor kiri) daripada nilai ambang c (Davino, Furno, dan Vistocco, 2014). Dalam banyak kasus, nilai ambang adalah 0. Model tersebut dapat diestimasi menggunakan Metode Maximum Likelihood (MLE) sebagai berikut.

$$\hat{\beta} = \arg \min_{\beta_\theta} \left\{ \prod_{i=1}^n (1 - F(\mathbf{x}_i^T \boldsymbol{\beta}))^{\Delta_i} f(y_i - \mathbf{x}_i^T \boldsymbol{\beta})^{1-\Delta_i} \right\}$$

dimana Δ_i menunjukkan indikator tersensor, $\Delta_i = 1$ jika observasi ke- i merupakan observasi tersensor, $\Delta_i = 0$ dan untuk lainnya. Namun fungsi F menghasilkan fungsi lain dalam *conditional mean* sehingga menyebabkan spesifikasi dari Gaussian estimasi maximum likelihood menjadi bias.

Powell (1986) mengamati bahwa kesamaan varians dari kuantil menjadi transformasi monoton menyebabkan model dari fungsi *conditional kuantil* dari variabel respon bergantung pada titik tersensor dan fungsi F yang independen. Secara resmi fungsi kuantil tersensor dari observasi respon y_i pada persamaan (2.23) dapat ditulis menjadi:

$$Q_{y_i}(\theta | x_i) = \max \{0, \mathbf{x}_i^T \boldsymbol{\beta} + F_\varepsilon^{-1}(\theta)\} \quad (2.24)$$

Parameter dari fungsi kuantil tersensor dapat diestimasi dengan mengganti persamaan:

$$\min_{\beta_\theta} \sum_{i=1}^n \rho_\theta(y_i - \mathbf{x}_i^T \boldsymbol{\beta})$$

dengan persamaan:

$$\min_{\beta_\theta} \sum_{i=1}^n \rho_\theta(y_i - \max \{0, \mathbf{x}_i^T \boldsymbol{\beta}\}) \quad (2.25)$$

dimana $\rho_\theta(\varepsilon)$ merupakan *check function* dengan definisi $\rho_\theta(\varepsilon) = \varepsilon(\theta - I[\varepsilon < 0])$ atau:

$$\rho_\theta(\varepsilon) = \begin{cases} \theta \varepsilon & , \varepsilon \geq 0 \\ (\theta - 1) \varepsilon & , \varepsilon < 0 \end{cases}$$

dimana $\theta \in (0,1)$, $\arg \min_{\beta_\theta}$ merupakan nilai yang meminimumkan nilai β dan $I(\cdot)$

merupakan fungsi indikator.

Penyelesaian kuantil ke θ , estimasi regresi kuantil tersensor adalah dengan menyelesaikan permasalahan minimasi fungsi *Least Absolute Deviations* (LAD) tersensor dibawah ini (Friederichs dan Hense, 2006):

$$\hat{\beta}_\theta = \arg \min_{\beta_\theta} \left[\sum_{i=1}^n \rho_\theta(y_i - \max\{0, \mathbf{x}_i^T \beta\}) \right] \quad (2.26)$$

Estimator regresi kuantil tersensor bisa diperoleh menggunakan solusi minimasi metode pemrograman linear. Buchinsky (1994) menawarkan penggunaan *Iterative Linear Programming Algorithm* (ILPA) yang melibatkan *Barrodale-Robert Algorithm* (BRA). Namun, metode ILPA memiliki kekurangan bahwa tidak ada kepastian konvergensi tercapai dan sekalipun tercapai, hal ini tidak menjamin solusi yang dihasilkan merupakan *local minima* dari permasalahan optimasi regresi kuantil tersensor. Selanjutnya Fitzenberger (1994) mengembangkan algoritma BRCENS sebagai adaptasi dari algoritma BRA dalam penjaminan konvergensi *local optima* yang dibangun atas karakteristik estimasi menggunakan *Interpolation Property*.

Simulasi studi yang dilakukan oleh Fitzenberger (1997) memperlihatkan bahwa algoritma BRCENS memberikan hasil yang lebih baik dari algoritma ILPA. Namun demikian seluruh algoritma tersebut memberikan performa yang kurang baik pada kondisi proporsi data tersensor yang besar dimana dari hasil simulasi pada kasus yang terdiri dari data tersensor sebanyak 50%, satu regressor, dan jumlah sampel yang kecil ($n=100$) frekuensi konvergen yang dihasilkan berada diantara 5% sampai dengan 37%. Dalam desain yang lain, hasil yang diperoleh mengalami perbaikan yaitu frekuensi konvergen berada diantara 30% sampai dengan 70%. Hal ini terjadi pada desain satu regressor, namun dalam kasus jumlah regressor yang lebih banyak dan jumlah sampel yang lebih besar, maka hasil yang diperoleh dapat lebih buruk. Dalam mengatasi permasalahan tersebut, dikembangkan tiga tahap algoritma yang dikenalkan oleh Chernozhukov dan Hong (2002) dimana tahapan yang digunakan lebih sederhana, robust, dan lebih dekat dengan titik sensing.

Tahapan algoritma ini adalah mengambil sub sampel dengan cara pemisahan dari probabilitas tersensor, dan melakukan dua kali estimasi menggunakan kuantil regresi. Hasil dari estimasi pertama adalah mendapatkan sub sampel yang sesuai, kemudian hasil dari estimasi yang kedua adalah untuk membuat estimasi yang efisien.

2.6 *Root Mean Square Error (RMSE)*

RMSE (*Root Mean Square Error*) merupakan salah satu statistik yang sering digunakan untuk mengevaluasi kebaikan performa model atau estimator. Statistik ini mengukur selisih antara nilai yang diprediksi oleh suatu model / estimator dengan nilai sebenarnya, yang disebut juga sebagai *error* atau residual. Persamaan berikut menunjukkan formulasi RMSE jika diasumsikan terdapat sebanyak n *error* model (Chai dan Draxler, 2014).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n e_i^2} \quad (2.27)$$

2.7 *Tinjauan Non Statistika*

Menurut Barber *et al.* (2008) komoditas rokok merupakan barang normal karena semakin tinggi harga barang tersebut maka jumlah permintaannya akan semakin berkurang. Akan tetapi pengaruh permintaan harga terhadap permintaan rokok diperkirakan kecil, artinya elastisitas permintaan karena harga (*price elasticity of demand*) kecil karena barang tersebut bersifat aditif (Hidayat dan Thabrany, 2010). Menurut Suranovic *et al.* (1999) sifat aditif rokok terlihat dari dua hal yaitu adanya efek menarik kembali perokok untuk mengkonsumsi rokok ketika berusaha untuk berhenti merokok serta seringkali dampak buruk merokok baru dirasakan pada akhir masa kehidupan seorang perokok.

Berdasarkan survey yang diteliti oleh World Healthy Organization (WHO) pada tahun 2011, Indonesia merupakan negara dengan jumlah perokok aktif ketiga terbesar di dunia setelah Cina dan India. Menurut data Riskesdas (2013) rata-rata proporsi perokok di Indonesia adalah 29,3 % atau rata-rata batang rokok yang dihisap per hari di Indonesia adalah 12,3 batang atau setara dengan satu

bungkus. Menurut hasil Susenas tahun 2015, penduduk berusia 15 tahun keatas yang mengkonsumsi rokok sebesar 22,57% berada di perkotaan dan 25,05% di pedesaan. Rata-rata jumlah batang rokok yang dihabiskan selama seminggu mencapai 76 batang di perkotaan dan 80 batang di pedesaan. Tingkat konsumsi rokok yang tinggi di masyarakat ini menunjukkan bahwa rokok merupakan produk yang permintaannya tinggi dan sudah menjadi salah satu kebutuhan masyarakat.

Salah satu indikator tingkat kesejahteraan masyarakat salah satunya bisa dilihat dengan memperhatikan pola konsumsi rumah tangga. Faktor-faktor yang memengaruhi tingkat konsumsi rumah tangga antara lain faktor ekonomi (pendapatan, kekayaan, tingkat bunga, dan perkiraan tentang masa depan), faktor demografi (jumlah dan komposisi penduduk) dan faktor nonekonomi (sosial budaya). Helmi (2016) melakukan penelitian terhadap faktor-faktor yang mempengaruhi perilaku merokok pada rumah tangga di kabupaten Lima Puluh Kota tahun 2013 memberikan hasil bahwa faktor-faktor yang memberikan pengaruh positif terhadap konsumsi rokok pada rumah tangga adalah pendapatan rumah tangga dan jumlah anggota rumah tangga. Sedangkan yang berpengaruh negatif adalah pendidikan kepala rumah tangga. Permana (2014) melakukan penelitian mengenai pengeluaran konsumsi rokok di kota Kediri tahun 2011 yang dibandingkan dengan umur, proporsi anggota rumah tangga dewasa, dan pendapatan rumah tangga. Cahyaningsih (2011) membuat pemodelan pengeluaran konsumsi rokok di Kalimantan Timur dan menjelaskan bahwa konsumsi rokok dipengaruhi secara signifikan oleh umur kepala rumah tangga, proporsi laki-laki, jumlah anggota rumah tangga, proporsi anggota rumah tangga dewasa, pendapatan rumah tangga, jenis kelamin kepala rumah tangga, pendidikan, sektor pekerjaan utama, serta keberadaan anak-anak dalam rumah tangga. Penjelasan mengenai penelitian sebelumnya akan disajikan secara ringkas pada Tabel 2.1 berikut ini.

Tabel 2.1 Penelitian Sebelumnya

| Nama Peneliti | Judul | Tahun | Variabel Penelitian |
|-----------------------|--|-------|---|
| Darma Putra Helmi | Analisis Faktor-Faktor Yang Mempengaruhi Perilaku Merokok Pada Rumah Tangga Di Kabupaten Lima Puluh Kota Tahun 2013. | 2016 | <ul style="list-style-type: none"> • Pendapatan rumah tangga • Jumlah anggota rumah tangga • Pendidikan kepala rumah tangga |
| Gilang Permana | Analisis regresi Tobit Pada Permasalahan Pengeluaran Konsumsi Rokok Kota Kediri Tahun 2011 | 2013 | <ul style="list-style-type: none"> • Umur kepala rumah tangga • Proporsi anggota rumah tangga dewasa • Pendapatan rumah tangga |
| Ariyanti Cahyaningsih | Pendekatan Tobit Model dan Double Hurdle Dalam Pemodelan Pengeluaran Konsumsi Rokok di Kalimantan Timur | 2011 | <ul style="list-style-type: none"> • Umur kepala rumah tangga • Proporsi laki-laki • Jumlah anggota rumah tangga • Proporsi anggota rumah tangga dewasa • Pendapatan rumah tangga • jenis kelamin kepala rumah tangga • pendidikan • sektor pekerjaan utama • keberadaan anak-anak dalam rumah tangga. |

Berdasarkan penelitian sebelumnya maka penelitian ini akan menggunakan variabel-variabel prediktor yang terdiri dari faktor-faktor sosial ekonomi yaitu pendapatan kepala rumah tangga, rata-rata pengeluaran per kapita, sektor pekerjaan kepala rumah tangga, serta faktor-faktor demografi yaitu tingkat pendidikan kepala rumah tangga, jumlah anggota rumah tangga, umur kepala rumah tangga, jenis kelamin kepala rumah tangga, serta daerah tempat tinggal meliputi wilayah perkotaan dan pedesaan.

BAB 3

METODE PENELITIAN

3.1 Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang dikeluarkan oleh World Bank melalui hasil Survei Sosial Ekonomi Rumah Tangga Indonesia (SUSETI) tahun 2011. SUSETI merupakan usaha kolaborasi antara Bank Dunia Kantor Jakarta (WBOJ), Biro Pusat Statistik (BPS), dan peneliti akademis yang bergabung dengan Jameel Poverty Action Lab (J-PAL) dengan cakupan wilayah penelitian meliputi 3 provinsi (Provinsi Sumatera Selatan, Provinsi Lampung dan Provinsi Jawa Tengah), 6 kabupaten, 63 kecamatan, 600 desa. Survei ini pertama kali dilakukan pada tahun 2008 dengan periode 4 tahun. Data tahun 2011 digunakan karena merupakan publikasi tahun terakhir. Unit sampel yang digunakan adalah rumah tangga dengan jumlah sampel sebanyak 5.507. Data yang digunakan adalah data pengeluaran rumah tangga untuk konsumsi rokok.

3.2 Variabel Penelitian

Variabel penelitian yang digunakan dibagi menjadi dua yaitu variabel respon dan variabel prediktor.

1. Variabel respon

Variabel respon yang digunakan adalah pengeluaran rumah tangga untuk konsumsi rokok selama satu bulan dalam ribuan rupiah.

$y_i = 0$ jika rumah tangga tidak mengeluarkan biaya untuk konsumsi rokok

$y_i = y_i^*$ jika rumah tangga mengeluarkan biaya untuk konsumsi rokok

2. Variabel Prediktor

- a. X_1 = Pendapatan rumah tangga selama satu bulan dalam ribuan rupiah ($\times 1000$).

- b. X_2 = Tingkat pendidikan kepala rumah tangga meliputi SD atau tidak tamat SD, SLTP dan SMU, serta Perguruan Tinggi. Kemudian variabel ini dirubah menjadi variabel *dummy* dengan struktur sebagai berikut:
 (0,0) untuk SD atau tidak tamat SD
 (1,0) untuk SLTP dan SMU
 (0,1) untuk Perguruan Tinggi
- c. X_3 = Jumlah anggota rumah tangga yaitu semua orang yang biasanya tinggal di dalam suatu rumah tangga (Jiwa).
- d. X_4 = Rata-rata pengeluaran per kapita yaitu total pengeluaran rumah tangga selama satu bulan dibagi dengan jumlah anggota rumah tangga dalam ribuan rupiah ($\times 1000$).
- e. X_5 = Umur kepala rumah tangga (Tahun)
- f. X_6 = Jenis kelamin kepala rumah tangga apakah laki-laki atau perempuan.
- g. X_7 = Sektor pekerjaan dari kepala rumah tangga meliputi:
- Sektor primer yaitu kepala rumah tangga yang bekerja pada sektor pertanian dan palawija, perkebunan / kehutanan, peternakan, perikanan, perburuan, dan pertambangan.
 - Sektor sekunder meliputi kepala rumah tangga yang bekerja pada sektor perindustrian / kerajinan, listrik, gas, air, dan bangunan / konstruksi.
 - Sektor tersier meliputi kepala rumah tangga yang bekerja pada sektor perdagangan, hotel, rumah makan, angkutan, pergudangan, komunikasi, keuangan, asuransi, perumahan, dan jasa kemasyarakatan.
 - Tidak bekerja meliputi kepala rumah tangga yang tidak bekerja.

Variabel tersebut akan dirubah menjadi variabel *dummy* dengan struktur sebagai berikut:

(0,0,0) untuk Sektor Primer

(0,0,1) untuk Sektor Sekunder

(0,1,0) untuk Sektor Tersier

(1,0,0) untuk Tidak Bekerja

- h. X_8 = Daerah tempat tinggal yaitu letak tempat tinggal suatu rumah tangga apakah di perkotaan atau pedesaan.

Struktur data yang digunakan ditampilkan pada Tabel 3.1.

Tabel 3.1 Struktur Data Untuk Analisis

| n | Y | X₁ | X_{2_1} | X_{2_2} | X₃ | X₄ | X₅ | X₆ | X_{7_1} | X_{7_2} | X_{7_3} | X₈ |
|----------|----------|----------------------|------------------------|------------------------|----------------------|----------------------|----------------------|----------------------|------------------------|------------------------|------------------------|----------------------|
| 1 | Y_1 | X_{11} | X_{211} | X_{221} | X_{31} | X_{41} | X_{51} | X_{61} | X_{711} | X_{721} | X_{731} | X_{81} |
| 2 | Y_2 | X_{12} | X_{212} | X_{222} | X_{32} | X_{42} | X_{52} | X_{62} | X_{712} | X_{722} | X_{732} | X_{82} |
| 3 | Y_3 | X_{13} | X_{213} | X_{223} | X_{33} | X_{43} | X_{53} | X_{63} | X_{713} | X_{723} | X_{733} | X_{83} |
| \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots |
| nt | Y_{nt} | $X_{1,nt}$ | $X_{2_1,nt}$ | $X_{2_2,n}$ | $X_{3,nt}$ | $X_{4,nt}$ | $X_{5,nt}$ | $X_{6,nt}$ | X_{7_1} | X_{7_2} | X_{7_3} | $X_{8,n}$ |

$n = 5.507$, $t=1$ =perkotaan, $t=2$ =pedesaan

3.3 Metode Penelitian

3.3.1 Estimasi Regresi Kuantil Tersensor

Kajian estimasi model regresi kuantil tersensor dilakukan dengan tahapan sebagai berikut:

1. Memformulasikan bentuk fungsional kuantil untuk variabel y_i .
2. Mengestimasi parameter model regresi kuantil tersensor dengan meminimumkan fungsi sebagai berikut.

$$Q_n(\beta; \theta) = \frac{1}{n} \sum_{i=1}^n \rho_{\theta}(y_i - \max\{0, \mathbf{x}_i^T \boldsymbol{\beta}\})$$

3. Menerapkan solusi minimasi metode pemrograman linear menggunakan tiga tahap algoritma yang dikenalkan oleh Chernozhukov dan Hong dengan tahapan sebagai berikut:

- a. Estimasi model probabilitas dalam sampel yang dituliskan pada persamaan $\Pr(y_i > 0|x_i) = F(x_i^T \gamma) + \varepsilon_i$ yang akan digunakan untuk memasukkan sub sampel $J_0 = \{i : x_i^T \hat{\gamma} > 1 - \theta + c\}$ dimana c adalah *trimming constant* diantara 0 dan 1.
- b. Mencari estimator awal $\hat{\beta}_\theta$ dengan memasukkan regresi kuantil biasa pada sampel J_0 . Gunakan estimator awal ini untuk memilih sampel $J_1 = \{x_i^T \hat{\beta}_\theta^0 > 0\}$ yang akan digunakan pada tahap selanjutnya.
- c. Estimasi model menggunakan regresi kuantil biasa pada sampel J_1 untuk mendapatkan estimasi $\hat{\beta}_\theta^1$.
- d. Jika diperlukan, ulangi step tahap 3 satu atau beberapa kali menggunakan sampel $J_I = \{x_i^T \hat{\beta}_\theta^I > 0\}$ dimana $I = 2, 3, \dots$ hingga diperoleh estimasi $\hat{\beta}_\theta$ yang efisien dalam arti bahwa pada step sebelumnya memiliki nilai konvergensi sebesar \sqrt{n} .

3.3.2 Perbandingan Metode Regresi Kuantil Tersensor

Perbandingan metode yang dilakukan adalah dengan membandingkan metode regresi kuantil tersensor dengan regresi kuantil dengan cara melihat nilai estimator masing-masing metode kemudian dilakukan perbandingan menggunakan nilai RMSE. Langkah-langkah perbandingan metode yang dilakukan adalah sebagai berikut.

1. Membangkitkan data tersensor dimana *error* mengikuti distribusi normal, ukuran sampel dibedakan menjadi 4 yaitu $n=100$, $n=500$, $n=1000$, dan $n=3000$.
2. Untuk setiap pembangkitan data, dilakukan estimasi parameter menggunakan metode regresi kuantil tersensor dengan regresi kuantil.
3. Menghitung nilai RMSE dari kedua metode tersebut kemudian melakukan perbandingan.

3.3.3 Penerapan Model Regresi Kuantil Tersensor

Langkah-langkah penerapan model regresi kuantil tersensor pada faktor-faktor yang mempengaruhi pengeluaran rumah tangga untuk konsumsi rokok adalah sebagai berikut:

1. Melakukan analisis deskriptif terhadap semua variabel yang digunakan dalam penelitian.
2. Membentuk model regresi tersensor untuk kuantil 0,1, 0,25, 0,5, 0,75, dan 0,9.
3. Mendapatkan estimasi parameter model regresi tersensor.
4. Menghitung nilai RMSE.
5. Interpretasi hasil.

(Halaman ini sengaja dikosongkan)

BAB 4

HASIL DAN PEMBAHASAN

4.1 Estimasi Regresi Kuantil Tersensor

Berikut ini adalah model regresi tersensor.

$$y_i = \begin{cases} \tau & y_i^* \leq \tau \\ \mathbf{x}_i^T \boldsymbol{\beta} + \varepsilon_i & y_i^* > \tau \end{cases} \quad (4.1)$$

$$y_i = \max(\tau, \mathbf{x}_i^T \boldsymbol{\beta} + \varepsilon_i) \quad (4.2)$$

Dengan τ adalah nilai titik sensor yang diketahui dan ditentukan $\tau = 0$ sehingga bentuk umum dari model regresi tersensor adalah:

$$y_i = \max(0, \mathbf{x}_i^T \boldsymbol{\beta}_0 + \varepsilon_i) \text{ atau } y_i = \max\{0, y_i^*\} \quad (4.3)$$

Dimana y_i^* merupakan variabel respon, vektor \mathbf{x}_i bersifat *observed* untuk masing-masing i , dan $i = 1, 2, \dots, n$. Sedangkan vektor $\boldsymbol{\beta}_0$ dan ε_i bersifat *unobserved*. Berdasarkan persamaan (2.23) diketahui bahwa:

$$y_i^* = \mathbf{x}_i^T \boldsymbol{\beta}_0 + \varepsilon_i, i = 1, 2, \dots, n$$

Definisi dari estimator *Least Absolut Deviation (LAD)* dalam model ini akan didasarkan pada kenyataan bahwa untuk variabel random skalar Z , pada fungsi $E[|Z - b| - |Z|]$ adalah dengan memilih b menjadi median dari distribusi Z (Powell, 1984). Karenanya jika median dari y_i adalah beberapa fungsi yang diketahui $m(x_i, \boldsymbol{\beta}_0)$ dari regressor dan parameter yang tidak diketahui, contoh perbandingan *conditional* median dapat didefinisikan dengan memilih $\hat{\boldsymbol{\beta}}_n$ pada fungsi $(1/n) \sum |y_i - m(x_i, \boldsymbol{\beta})|$ adalah dengan meminimalkan nilai $\boldsymbol{\beta} = \hat{\boldsymbol{\beta}}_n$. Bentuk sederhana dari fungsi median dari y_i adalah $m(x_i, \boldsymbol{\beta}_0) = \max\{0, \mathbf{x}_i^T \boldsymbol{\beta}_0\}$. Sehingga estimasi *Least Absolut Deviation (LAD)* tersensor untuk $\hat{\boldsymbol{\beta}}_n$ adalah dengan meminimalkan nilai $\boldsymbol{\beta}$ seperti dalam persamaan berikut:

$$S_n(\boldsymbol{\beta}) = \frac{1}{n} \sum_{i=1}^n |y_i - \max\{0, \mathbf{x}_i^T \boldsymbol{\beta}_0\}| \quad (4.4)$$

Untuk keseluruhan β dalam beberapa ruang parameter B . Definisi ini didasarkan pada hubungan sederhana antara median dari variabel dependen tersensor dan regresor dengan vektor parameter.

Selanjutnya estimasi parameter β pada regresi kuantil menggunakan *Least Absolut Deviation (LAD)* dilakukan dengan meminimumkan kuadrat error yaitu:

$$\hat{\beta}_{\theta} = \arg \min_{\beta_{\theta}} \left[\sum_{i=1}^n \rho_{\theta}(y_i - \mathbf{x}_i^T \beta) \right] \quad (4.5)$$

dimana $\rho_{\theta}(\varepsilon)$ merupakan *loss function* dengan definisi $\rho_{\theta}(\varepsilon) = \varepsilon(\theta - I[\varepsilon < 0])$ atau:

$$\rho_{\theta}(\varepsilon) = \begin{cases} \theta \varepsilon & , \varepsilon \geq 0 \\ (\theta - 1) \varepsilon & , \varepsilon < 0 \end{cases}$$

dimana $\theta \in (0,1)$, $\arg \min_{\beta_{\theta}}$ merupakan nilai yang meminimumkan nilai β dan

$I(\cdot)$ merupakan fungsi indikator. Penjelasan mengenai *loss function* terdapat pada Lampiran 1.

Tahap awal dalam mendefinisikan estimator regresi kuantil tersensor adalah menentukan bentuk fungsional dari kuantil pada variabel dependen y_i .

$F_Y^{-1}(\theta | x_i, \beta_0)$ menunjukkan kuantil ke- θ dari y_i untuk $\theta \in (0,1)$. Sehingga fungsi kuantil bersyarat dari y_i adalah nilai kanan dari persamaan (4.3) ketika ε_i digantikan dengan kuantil ke- θ .

$$F_Y^{-1}(\theta | x_i, \beta_0) = \max \{0, \mathbf{x}_i^T \beta_0 + F^{-1}(\theta)\} \quad (4.6)$$

Median dari ε_i adalah $F^{-1}\left(\frac{1}{2}\right) = 0$ sehingga komponen pertama dalam vektor parameter β_0 adalah sebuah intercept, maka kuantil ke- θ dari y_i dapat ditulis sebagai berikut:

$$F_Y^{-1}(\theta | x_i, \beta_0) = \max \{0, \mathbf{x}_i^T \beta_0(\theta)\} \quad (4.7)$$

Dimana:

$$\beta_0(\theta) = \beta_0 + F^{-1}(\theta) \cdot \mathbf{e}^1 \quad \mathbf{e}^1 = (1, 0, \dots, 0)^T \quad (4.8)$$

Dengan mendapatkan informasi variabel tersensor, maka parameter dari fungsi kuantil tersensor dapat diestimasi dengan mengganti persamaan:

$$\min_{\beta_\theta} \sum_{i=1}^n \rho_\theta(y_i - \mathbf{x}_i^T \boldsymbol{\beta}) \quad (4.9)$$

dengan persamaan:

$$\min_{\beta_\theta} \sum_{i=1}^n \rho_\theta(y_i - \max\{0, \mathbf{x}_i^T \boldsymbol{\beta}\}) \quad (4.10)$$

Untuk mendapatkan estimator $\beta_0(\theta)$ pada persamaan (4.6) untuk nilai tertentu θ adalah sejalan dengan persamaan (4.4) dengan menggunakan *loss function*. Sehingga estimator regresi kuantil tersensor $\hat{\beta}_n(\theta)$ dari $\beta_0(\theta)$ didefinisikan meminimalkan nilai β berikut ini.

$$Q_n(\beta; \theta) = \frac{1}{n} \sum_{i=1}^n \rho_\theta(y_i - \max\{0, \mathbf{x}_i^T \boldsymbol{\beta}\}) \quad (4.11)$$

Karena ketidakterkaitan dalam fungsi objektif, maka permasalahan optimasi ini membutuhkan penyelesaian secara komputasi menggunakan algoritma optimasi. Chernozhukov dan Hong (2002) memperkenalkan tiga tahap algoritma yang dapat digunakan pada proporsi data tersensor yang besar.

Menerapkan solusi minimasi metode pemrograman linear menggunakan tiga tahap algoritma yang dikenalkan oleh Chernozhukov dan Hong dengan tahapan sebagai berikut:

- a. Estimasi model menggunakan metode regresi parametrik untuk data biner yang dituliskan pada persamaan: $\Pr(y_i > 0 | x_i) = F(x_i^T \gamma) + \varepsilon_i$.

Selanjutnya model tersebut akan digunakan untuk memasukkan sub sampel $J_0 = \{i : x_i^T \hat{\gamma} > 1 - \theta + c\}$ dimana c adalah *trimming constant* diantara 0 dan 1. Tujuan dari step 1 ini adalah untuk memilih subset / bagian dari observasi dimana $\Pr(y_i > 0 | x_i) > 1 - \theta$, yaitu dimana garis kuantil $\mathbf{x}_i^T \boldsymbol{\beta}_\theta$ berada diatas titik tersensor. Pemilihan *trimming constant* dapat menggunakan perbandingan ukuran pada sampel yang terambil pada $J_c = \{i : x_i^T \hat{\gamma} > 1 - \theta + c\}$ untuk nilai $c = 0$ dan nilai

lainnya. Pemilihan $c =$ kuantil ke- q dan mengikuti persamaan $\frac{\#J_c}{\#J_0} = (1 - q)\%$ dimana nilai q ditentukan 5% dan 10% . Aturan lain yang dapat digunakan untuk menentukan *trimming constant* adalah dengan menentukan langsung yaitu 0,01, 0,02, 0,03, dan seterusnya (Gustavsen *et al.*2008).

- b. Mencari estimator awal $\hat{\beta}_\theta$ dengan menjalankan standar regresi kuantil

$$\text{pada sampel } J_0 : \min_{i \in J_0} \sum \rho_\theta(y_i - x_i^T \beta)$$

Gunakan estimator awal ini untuk memilih sampel $J_1 = \{x_i^T \hat{\beta}_\theta^0 > 0\}$ yang akan digunakan pada tahap selanjutnya.

- c. Estimasi model menggunakan standar regresi kuantil dengan

$$\text{mengganti sampel } J_0 \text{ menjadi } J_1 : \min_{i \in J_1} \sum \rho_\theta(y_i - x_i^T \beta) \text{ untuk}$$

mendapatkan estimasi $\hat{\beta}_\theta^1$.

- d. Jika diperlukan ulangi step tahap c beberapa kali menggunakan sampel

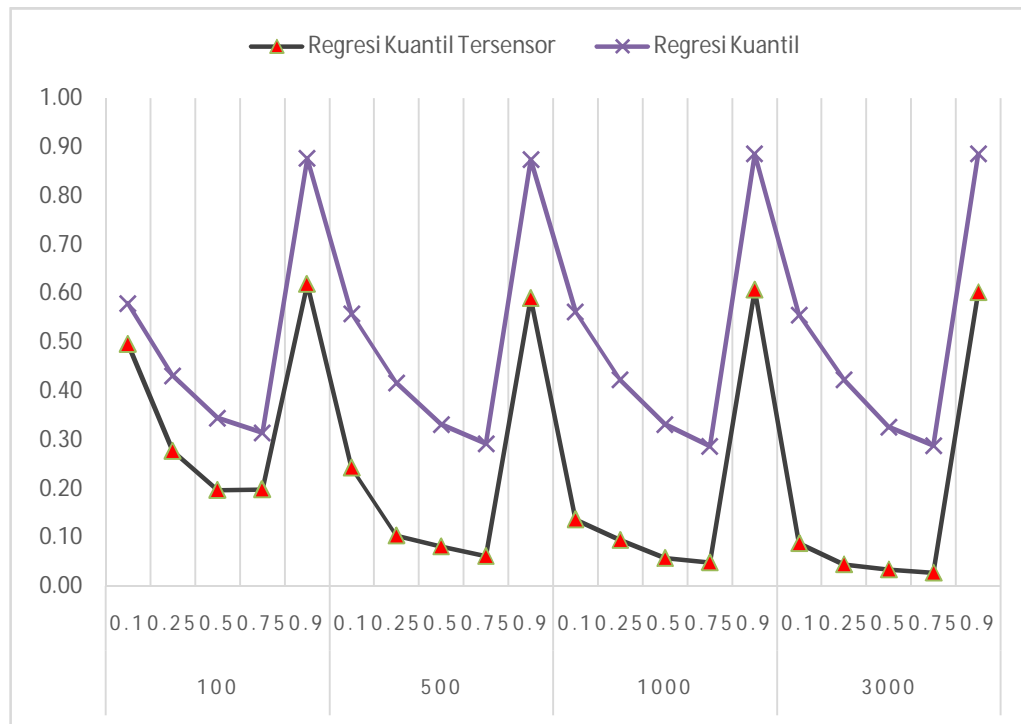
$J_I = \{x_i^T \hat{\beta}_\theta > 0\}$ dimana $I = 2, 3, \dots$ hingga diperoleh estimasi $\hat{\beta}_\theta$ yang efisien dalam arti bahwa pada step sebelumnya memiliki nilai konvergensi sebesar \sqrt{n} .

4.2 Perbandingan Metode Regresi Kuantil Tersensor

Model regresi kuantil tersensor merupakan pengembangan dari metode regresi kuantil. Oleh karena itu, akan dilakukan perbandingan metode antara regresi kuantil tersensor dan regresi kuantil. Metode perbandingan yang digunakan adalah dengan membangkitkan data menggunakan ukuran sampel berbeda dengan jumlah variabel prediktor yang sama, Ukuran pembanding yang digunakan adalah RMSE, dimana semakin kecil nilai RMSE maka performa estimator akan semakin baik. Berikut akan disajikan hasil perbandingan metode regresi kuantil tersensor dan regresi kuantil pada Tabel 4.1 serta akan divisualisasikan pada Gambar 4.1.

Tabel 4.1 RMSE Intersep dari Estimator Regresi Kuantil Tersensor dan Regresi Kuantil

| Metode | n | Kuantil | | | | |
|---------------------------|------|---------|------|------|------|------|
| | | 0,1 | 0,25 | 0,5 | 0,75 | 0,9 |
| Regresi Kuantil Tersensor | 100 | 0,50 | 0,28 | 0,20 | 0,20 | 0,62 |
| Regresi Kuantil | | 0,58 | 0,43 | 0,34 | 0,31 | 0,88 |
| Regresi Kuantil Tersensor | 500 | 0,24 | 0,10 | 0,08 | 0,06 | 0,59 |
| Regresi Kuantil | | 0,56 | 0,42 | 0,33 | 0,29 | 0,87 |
| Regresi Kuantil Tersensor | 1000 | 0,14 | 0,09 | 0,06 | 0,05 | 0,61 |
| Regresi Kuantil | | 0,56 | 0,42 | 0,33 | 0,29 | 0,89 |
| Regresi Kuantil Tersensor | 3000 | 0,09 | 0,04 | 0,03 | 0,03 | 0,60 |
| Regresi Kuantil | | 0,56 | 0,42 | 0,33 | 0,29 | 0,89 |



Gambar 4.1 RMSE Intersep dari Estimator Regresi Kuantil Tersensor dan Regresi Kuantil

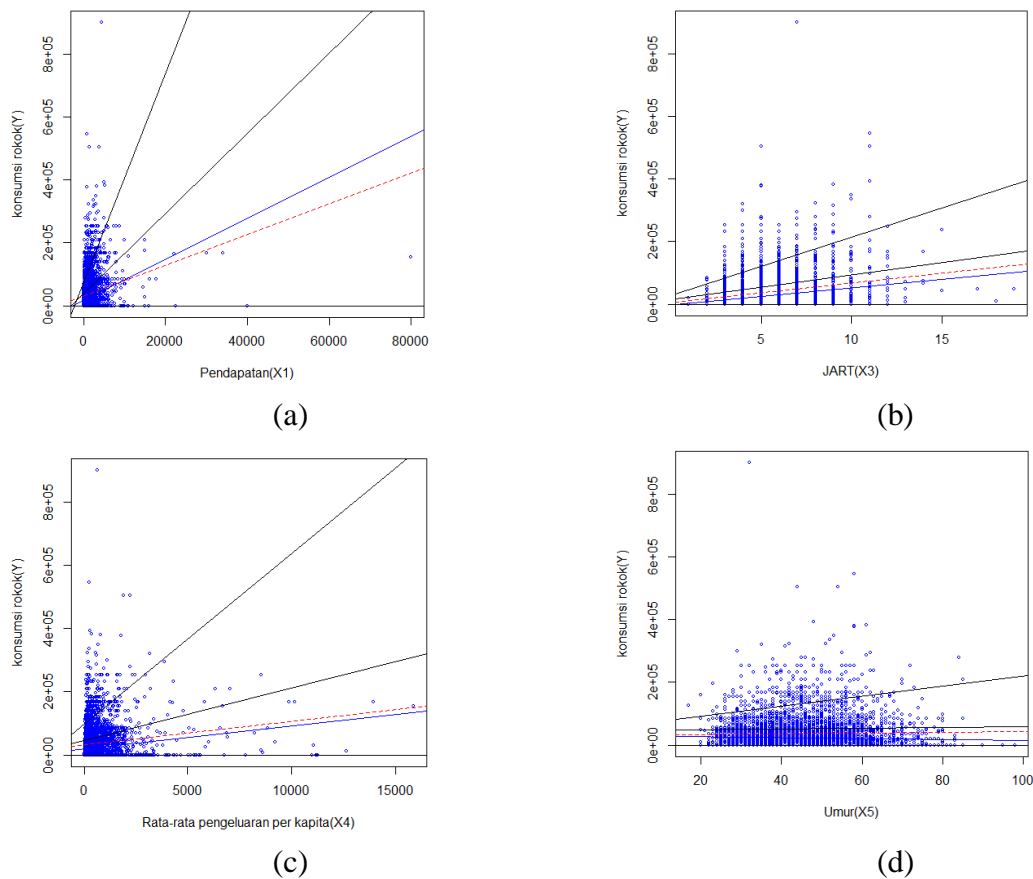
Gambar 4.1 merupakan hasil simulasi ketiga metode menggunakan jumlah sampel yang berbeda yaitu 100, 500, 1000, dan 3000. Jumlah variabel prediktor yang digunakan adalah 3 variabel dengan *error* mengikuti distribusi normal. Dari Gambar 4.6 dapat diperoleh informasi bahwa pada kondisi terdapat data tersensor maka performa estimator regresi kuantil tersensor cukup baik pada setiap jumlah sampel yang berbeda. Nilai RMSE estimator regresi kuantil tersensor lebih kecil

daripada regresi kuantil, atau dengan kata lain performa estimator regresi kuantil tersensor lebih baik daripada regresi kuantil.

4.3 Penerapan Model Regresi Kuantil Tersensor

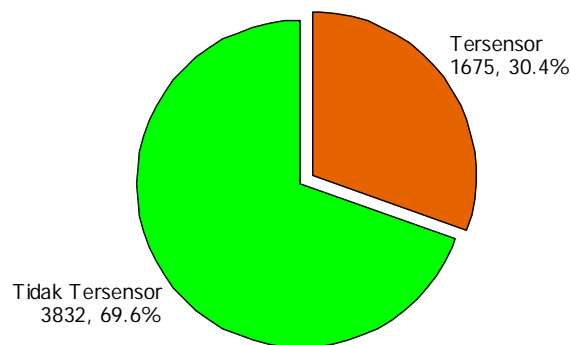
4.3.1 Deskriptif Data Penelitian

Langkah pertama dalam melakukan penerapan model regresi kuantil tersensor adalah dengan melakukan analisis deskriptif dari seluruh variabel yang digunakan. Berikut akan ditampilkan *scatter plot* diantara variabel respon dengan masing-masing variabel prediktor yang bersifat kontinu untuk mengetahui sebaran data yang terjadi diantara variabel-variabel tersebut.



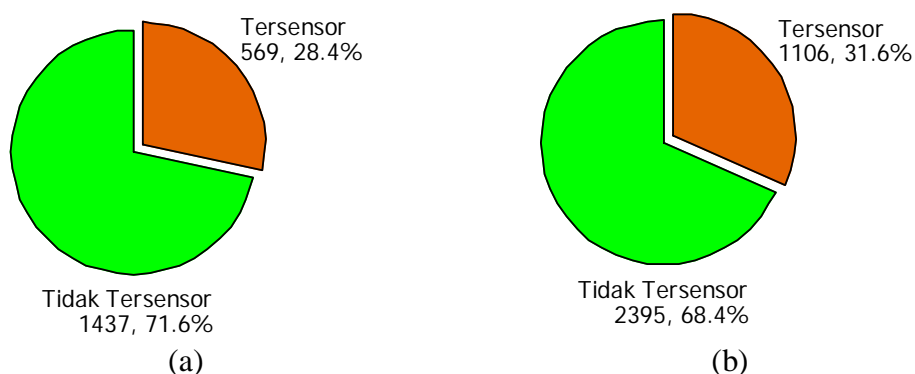
Gambar 4.2 *Scatter Plot* berdasarkan kuantil : (a) variabel konsumsi rokok dengan pendapatan, (b) variabel konsumsi rokok dengan jumlah anggota rumah tangga (JART), (c) variabel konsumsi rokok dengan rata-rata pengeluaran per kapita, (d) variabel konsumsi rokok dengan umur kepala rumah tangga.

Pada Gambar 4.2 (a) dapat diperoleh informasi bahwa rumah tangga yang mengeluarkan biaya untuk konsumsi rokok mayoritas dikeluarkan oleh rumah tangga yang memiliki penghasilan yang rendah. Sedangkan jika melihat dari sebaran data, maka data banyak berkumpul pada kuantil diatas median, hal ini menunjukkan bahwa lebih dari setengah rumah tangga pada kelompok rumah tangga dengan penghasilan yang rendah memiliki pengeluaran untuk konsumsi rokok yang tinggi. Pada gambar (b) dapat diperoleh informasi bahwa banyaknya jumlah anggota rumah tangga tidak memberikan pengaruh yang berarti terhadap banyaknya pengeluaran rumah tangga untuk konsumsi rokok. Hal ini dapat dilihat dari sebaran data yang terbentuk menyebar secara acak. Hal yang sama terjadi pada gambar (d) yang menunjukkan sebaran data yang acak atau tidak membuat pola tertentu, dimana data menyebar hampir di setiap garis-garis kuantil sehingga dapat dikatakan bahwa hampir semua kelompok umur kepala rumah tangga memberikan pengaruh yang besar terhadap pengeluaran rumah tangga untuk konsumsi rokok. Selanjutnya pada gambar (c) dapat diperoleh informasi bahwa rata-rata pengeluaran per kapita berkumpul di nilai yang rendah, dapat juga digambarkan bahwa tidak terjadi hubungan yang positif dengan pengeluaran rumah tangga untuk konsumsi rokok dimana kenaikan rata-rata pengeluaran rumah tangga tidak diikuti dengan kenaikan pengeluaran rumah tangga untuk konsumsi rokok. Jika dilihat dari sebaran data, maka data banyak berkumpul pada kuantil diatas median sehingga dapat dikatakan bahwa mayoritas kelompok rumah tangga memiliki pengeluaran untuk konsumsi rokok yang besar. Secara keseluruhan sebaran data yang terjadi diantara variabel respon dengan masing-masing variabel prediktor yang bersifat kontinu memiliki sebaran data yang tidak seragam di setiap kuantil sehingga perlu digali kembali bagaimana hubungan antara variabel pengeluaran rumah tangga untuk konsumsi rokok dengan masing-masing variabel prediktor di masing-masing kuantil. Selanjutnya statistik deskriptif dari data yang digunakan dalam penelitian ini dapat ditampilkan pada Gambar 4.3, Gambar 4.4, Tabel 4.2, dan Tabel 4.3



Gambar 4.3 Diagram Pie Kategori dalam Variabel Respon

Dalam penelitian ini ditentukan titik tersensor yang digunakan adalah 0, maka dapat dikatakan bahwa nilai 0 pada variabel respon yaitu pengeluaran rumah tangga untuk konsumsi rokok disebut sebagai data tersensor dan untuk nilai lainnya disebut sebagai data tidak tersensor. Pada Gambar 4.3 dapat dilihat bahwa banyaknya data tersensor adalah sebesar 1.675 rumah tangga atau dapat dikatakan bahwa sebesar 30,4% rumah tangga tidak memiliki pengeluaran untuk konsumsi rokok. Sedangkan data tidak tersensor sebanyak 3.832 rumah tangga atau sebesar 69,6% rumah tangga yang memiliki pengeluaran untuk konsumsi rokok. Selanjutnya proporsi mengenai kategori variabel respon untuk wilayah perkotaan dan pedesaan akan ditunjukkan pada Gambar 4.4.



Gambar 4.4 Diagram Pie Kategori Variabel Respon (a) Perkotaan, (b) Pedesaan

Pada Gambar 4.4 dapat diperoleh informasi bahwa di wilayah perkotaan terdapat 28,4% rumah tangga yang tidak memiliki pengeluaran untuk konsumsi rokok, sedangkan sebesar 71,6% rumah tangga memiliki pengeluaran untuk

konsumsi rokok. Pada wilayah pedesaan terdapat 31,6% rumah tangga yang tidak memiliki pengeluaran untuk konsumsi rokok, sedangkan sebesar 68,4% rumah tangga memiliki pengeluaran untuk konsumsi rokok. Selanjutnya akan dijelaskan mengenai statistik deskriptif untuk data tidak tersensor.

Tabel 4.2 Statistik Deskriptif Pengeluaran Rumah Tangga untuk Konsumsi Rokok (Ribuan Rupiah)

| Statistik | Wilayah | | |
|-----------------|--------------------------------------|-----------|----------|
| | Gabungan (Pedesaan dan Perkotaan) | Perkotaan | Pedesaan |
| Minimum | 600 | 600 | 1.000 |
| Maksimum | 900.000 | 900.000 | 504.000 |
| Rata – rata | 53.219 | 60.871 | 48.628 |
| Standar Deviasi | 50.708 | 59.348 | 44.111 |

Berdasarkan Tabel 4.2 dapat dilihat bahwa untuk wilayah gabungan (perkotaan dan pedesaan) pengeluaran rumah tangga untuk konsumsi rokok adalah sebesar Rp 900.000,-. Sedangkan jika dibandingkan antara wilayah perkotaan dan pedesaan dapat terlihat bahwa di wilayah perkotaan pengeluaran rumah tangga untuk konsumsi rokok lebih tinggi di bandingkan di wilayah pedesaan. Hal ini sejalan dengan rata-rata dan standar deviasi pengeluaran rumah tangga untuk konsumsi rokok di wilayah perkotaan yang lebih besar daripada di wilayah pedesaan.

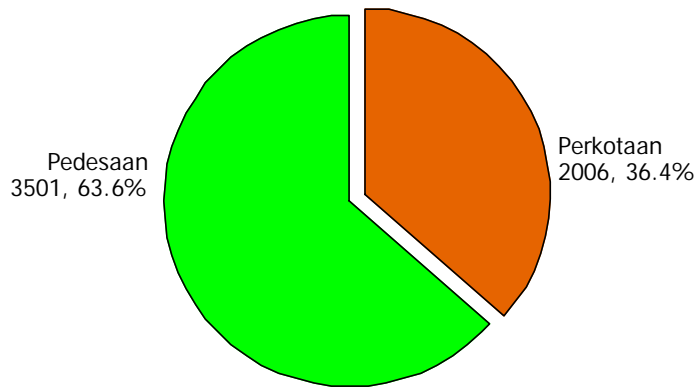
Tabel 4.3 Statistik Deskriptif Variabel Prediktor yang Bersifat Kontinu

| Statistik | Variabel | | | | |
|-----------------|-----------|-----------------------------|--------------------------|-----------------------------|---------------------------|
| | Wilayah | X ₁ (Ribu Rp) | X ₃ (Jiwa) | X ₄ (Ribu Rp) | X ₅ (Tahun) |
| Minimum | Gabungan | 100 | 1 | 24,4 | 17 |
| | Perkotaan | 100 | 2 | 0,57 | 20 |
| | Pedesaan | 100 | 1 | 0,29 | 17 |
| Maksimum | Gabungan | 80.000 | 19 | 15.879,8 | 98 |
| | Perkotaan | 34.000 | 19 | 76,58 | 85 |
| | Pedesaan | 80.000 | 18 | 75,94 | 98 |
| Rata – rata | Gabungan | 1.319 | 5,0632 | 512 | 44,253 |
| | Perkotaan | 1.742,8 | 5,4422 | 21,673 | 45,012 |
| | Pedesaan | 1.076,2 | 4,846 | 25,825 | 43,818 |
| Standar Deviasi | Gabungan | 1.984 | 1,762 | 847,7 | 11,248 |
| | Perkotaan | 2.086,9 | 1,9396 | 12,745 | 10,928 |
| | Pedesaan | 1.880,3 | 1,6123 | 14,388 | 11,406 |

Berdasarkan Tabel 4.3 dapat dilihat bahwa pada variabel pendapatan (X_1) nilai rata-rata dan standar deviasi gabungan di wilayah perkotaan dan pedesaan masing-masing adalah 1.742,8 dan 1.076. Jika dilihat di wilayah perkotaan dan pedesaan maka rata-rata dan standar deviasi di wilayah perkotaan lebih besar daripada wilayah pedesaan. Jika dilihat pendapatan maka di wilayah perkotaan memiliki pendapatan yang jauh lebih besar dibandingkan wilayah pedesaan. Hal ini disebabkan karena di wilayah perkotaan memiliki lapangan pekerjaan yang lebih luas dan meliputi banyak sektor dibandingkan dengan wilayah pedesaan.

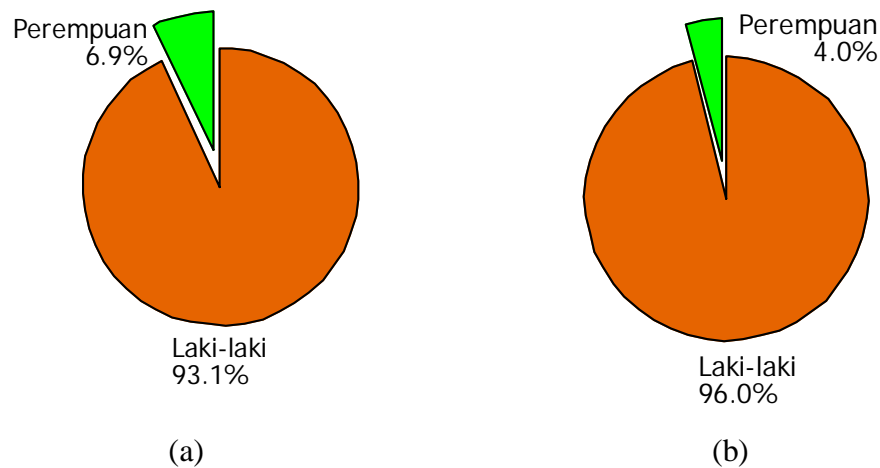
Selanjutnya untuk variabel rata-rata pengeluaran per kapita (X_4) wilayah pedesaan memiliki rata-rata pengeluaran per kapita yang lebih tinggi dibandingkan wilayah perkotaan. Namun jika dilihat dari nilai maksimum maka di wilayah perkotaan memiliki rata-rata pengeluaran per kapita yang lebih tinggi dibandingkan wilayah pedesaan. Hal ini disebabkan karena di perkotaan banyak memiliki pengeluaran atau biaya hidup yang lebih tinggi dibandingkan dengan di pedesaan.

Menurut aspek demografis, jumlah anggota keluarga (X_3) di wilayah perkotaan memiliki jumlah anggota rumah tangga yang lebih besar daripada di wilayah pedesaan, hal ini sejalan dengan nilai rata-rata dan standar deviasi wilayah perkotaan yang juga lebih tinggi dibandingkan di pedesaan. Untuk variabel umur kepala rumah tangga (X_5) di wilayah Perkotaan, umur kepala rumah tangga minimal adalah pada usia 20 tahun, sedangkan di wilayah pedesaan adalah 17 tahun. Nilai maksimum umur kepala rumah tangga di pedesaan adalah 98 tahun sehingga jauh lebih tinggi dibandingkan di wilayah perkotaan yang hanya 85 tahun. Wilayah perkotaan memiliki rata-rata yang lebih tinggi dibandingkan dengan wilayah pedesaan.



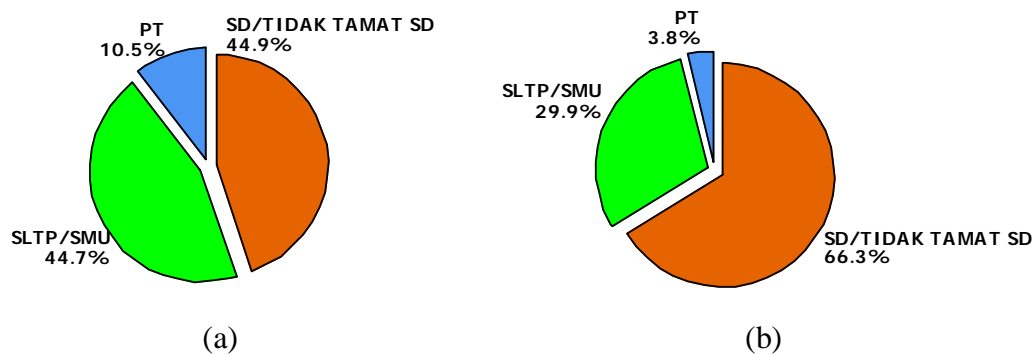
Gambar 4.5 Diagram Pie Wilayah Tempat Tinggal

Pada Gambar 4.5 dapat diperoleh informasi mengenai wilayah tempat tinggal responden dimana terdapat dua wilayah yaitu wilayah pedesaan dan wilayah perkotaan. Responden terbanyak dalam penelitian ini berasal dari wilayah pedesaan yaitu sebanyak 3.501 responden atau 63,6%, sedangkan responden yang berasal dari wilayah perkotaan adalah sebanyak 2006 responden atau hanya 36,4%.



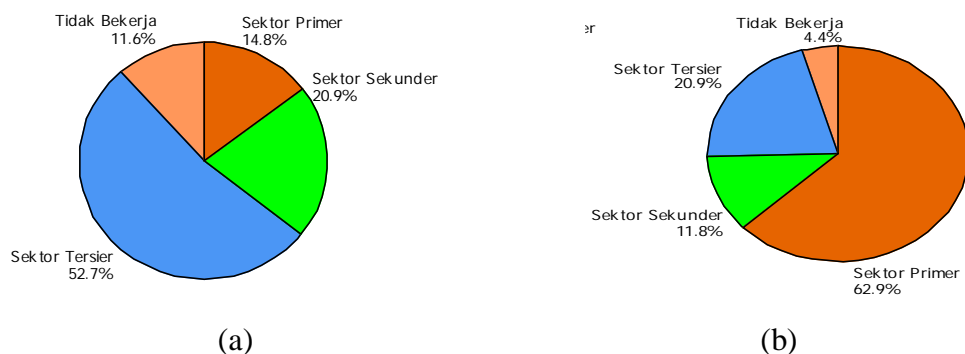
Gambar 4.6 Diagram Pie Jenis Kelamin (a) Perkotaan, (b) Pedesaan

Pada Gambar 4.6 dapat diperoleh informasi mengenai jenis kelamin kepala rumah tangga di wilayah perkotaan dan pedesaan. Dari kedua wilayah sebagian besar kepala keluarga didominasi oleh jenis kelamin laki-laki, sedangkan hanya sebagian kecil saja rumah tangga yang kepala keluarganya berjenis kelamin perempuan.



Gambar 4.7 Diagram Pie Pendidikan (a) Perkotaan, (b) Pedesaan

Pada Gambar 4.7 dapat diperoleh informasi mengenai pendidikan kepala rumah tangga baik di wilayah perkotaan dan pedesaan yang mayoritas memiliki pendidikan SD atau tidak tamat SD, namun persentase responden yang memiliki pendidikan SD atau tidak tamat SD di pedesaan lebih besar daripada di wilayah perkotaan. Pada urutan kedua pendidikan SLTP atau SMU di wilayah perkotaan jauh lebih tinggi dibandingkan wilayah pedesaan. Hal ini terjadi juga pada tingkat pendidikan Perguruan Tinggi di perkotaan yang lebih tinggi dibandingkan dengan wilayah pedesaan.



Gambar 4.8 Diagram Pie Sektor Pekerjaan (a) Perkotaan, (b) Pedesaan

Pada Gambar 4.8 dapat diperoleh informasi mengenai sektor pekerjaan di perkotaan dan pedesaan. Dari variabel sektor pekerjaan terdapat perbedaan yang besar diantara sektor pekerjaan kepala rumah tangga di wilayah perkotaan dan pedesaan. Pada wilayah perkotaan, mayoritas kepala rumah tangga bekerja pada sektor tersier yaitu sektor perdagangan, hotel, rumah makan, angkutan,

pergudangan, komunikasi, keuangan, asuransi, perumahan, dan jasa kemasyarakatan. Pada urutan selanjutnya adalah rumah tangga yang bekerja pada sektor sekunder, primer, dan tidak bekerja. Berbeda dengan wilayah perkotaan, di wilayah pedesaan mayoritas kepala rumah tangga memiliki pekerjaan pada sektor primer yaitu sektor pertanian dan palawija, perkebunan / kehutanan, peternakan, perikanan, perburuan, dan pertambangan. Pada urutan selanjutnya adalah rumah tangga yang bekerja pada sektor tersier, sekunder, dan tidak bekerja.

4.3.2 Hasil Estimasi Parameter Regresi Kuantil Tersensor

Estimasi parameter regresi kuantil tersensor menggunakan data pengeluaran rumah tangga untuk konsumsi rokok. Dengan melihat analisis deskriptif bahwa terdapat perbedaan karakteristik diantara perkotaan dan pedesaan, sehingga estimasi parameter akan dilakukan pada 3 kondisi, pertama akan dilakukan pemodelan pada data pengeluaran rumah tangga untuk konsumsi rokok secara keseluruhan, tidak membedakan wilayah perkotaan dan pedesaan dengan jumlah sampel 5.507. Kondisi kedua adalah dilakukan pemodelan pada data pengeluaran rumah tangga untuk konsumsi rokok di wilayah perkotaan dengan jumlah sampel 2.006. Kondisi terakhir adalah pemodelan untuk data pengeluaran rumah tangga untuk konsumsi rokok di wilayah perkotaan dengan jumlah sampel 3.501. Estimasi parameter regresi kuantil tersensor pada model gabungan (perkotaan dan pedesaan) dilakukan untuk lima titik kuantil yaitu 0,1, 0,25, 0,5, 0,75, 0,9. Sedangkan pada model untuk wilayah perkotaan dan pedesaan akan dilakukan untuk empat kuantil yaitu 0,25, 0,5, 0,75, 0,9. Sedangkan sebagai bahan perbandingan akan dilakukan analisis menggunakan metode regresi kuantil pada data pengeluaran rumah tangga untuk konsumsi rokok. Analisis regresi kuantil dilakukan pada kondisi yang sama dengan analisis regresi kuantil tersensor. Adapun hasil estimasi parameter menggunakan metode regresi kuantil tersensor akan disajikan pada Tabel 4.4, 4.5, dan 4.6. Sedangkan hasil estimasi parameter menggunakan metode regresi kuantil akan disajikan pada Tabel 4.7, 4.8, dan 4.9.

Tabel 4.4 Estimasi Parameter Regresi Kuantil Tersensor di Wilayah Gabungan

| Estimasi Parameter | Estimator Kuantil | | | | |
|-----------------------|----------------------|---------|--------|--------|-------|
| | 0,1 | 0,25 | 0,5 | 0,75 | 0,9 |
| β_0 | -80,02 | -4,51 | 17,61 | 23,50 | 33,1 |
| β_1 | 0,0017 | 0,001 | 0,005 | 0,01 | 0,01 |
| β_{2_1} | -32,32 | -12,33 | -27,58 | -16,89 | -8,81 |
| β_{2_2} | -12,4 | -1,67 | -2,02 | -2,39 | -1,49 |
| β_3 | 6,76 | 2,27 | 5,09 | 6,1 | 10,73 |
| β_4 | -0,0049 | -0,0009 | 0,004 | 0,009 | 0,02 |
| β_5 | 0,11 | -0,06 | -0,48 | -0,24 | -0,37 |
| β_6 | -14,1 | -20,48 | -15,77 | -16,96 | 4,65 |
| β_{7_1} | 27,2 | 0,82 | 1,22 | 1,8 | -6,9 |
| β_{7_2} | 1,63 | -0,33 | -2,009 | 0,94 | -0,08 |
| β_{7_3} | 28,62 | 0,01 | -3,85 | -3,35 | -9,16 |
| β_8 | 2,22 | -1,99 | -4,55 | -5,65 | -9,89 |

Dari Tabel 4.4 dapat diperoleh informasi mengenai efek tiap parameter pada masing-masing kuantil terhadap pengeluaran rumah tangga untuk konsumsi rokok. Dari hasil estimasi tiap parameter menunjukkan bahwa pengeluaran rumah tangga untuk konsumsi rokok bervariasi antar kuantil, hal ini menunjukkan bahwa penggunaan model regresi kuantil tersensor sudah tepat digunakan dalam pemodelan data tersebut. Besarnya estimator pada beberapa variabel mengalami kecenderungan semakin meningkat seiring pertambahan nilai kuantil. Hal ini menunjukkan bahwa pengaruh variabel tersebut semakin besar pada pengeluaran rumah tangga untuk konsumsi rokok yang semakin tinggi. Variabel yang memberikan pengaruh yang besar untuk konsumsi rokok yang tinggi adalah variabel pendapatan (X_1), tingkat pendidikan kepala rumah tangga menengah-rendah (X_{2_1}), tingkat pendidikan kepala rumah tangga tinggi-rendah (X_{2_2}), jumlah anggota rumah tangga (X_3), rata-rata pengeluaran per kapita (X_4), dan jenis kelamin kepala rumah tangga (X_6). Selain itu terdapat beberapa variabel yang memiliki kecenderungan nilai estimator yang semakin kecil di setiap kenaikan kuantil, hal ini menunjukkan bahwa pada kenaikan pengeluaran rumah tangga untuk konsumsi rokok maka pengaruh dari variabel tersebut akan semakin kecil. Variabel yang memberikan pengaruh yang kecil pada kenaikan pengeluaran konsumsi rokok adalah variabel umur (X_5), sektor pekerjaan di sektor sekunder-

primer (X_{7_1}), sektor pekerjaan di sektor tersier-primer (X_{7_2}), sektor pekerjaan di sektor tidak bekerja-primer (X_{7_3}), dan wilayah tempat tinggal (X_8).

Tabel 4.5 Estimasi Parameter Regresi Kuantil Tersensor di Wilayah Perkotaan

| Estimasi Parameter | Estimator Kuantil | | | |
|-----------------------|----------------------|--------|--------|--------|
| | 0,25 | 0,5 | 0,75 | 0,9 |
| β_0 | -36,8 | -3,34 | 6,78 | 14,93 |
| β_1 | 0,0041 | 0,005 | 0,01 | 0,02 |
| β_{2_1} | -21,91 | -14,36 | -21,27 | -3,74 |
| β_{2_2} | -3,96 | -4,24 | -7,61 | -8,25 |
| β_3 | 4,41 | 3,44 | 5,1 | 8,54 |
| β_4 | 0,82 | 0,59 | 0,46 | 0,67 |
| β_5 | -0,12 | -0,05 | 0,05 | 0,1 |
| β_6 | -30,22 | -22,02 | -23,94 | -14,18 |
| β_{7_1} | -1,28 | 2,58 | 2,44 | -15,59 |
| β_{7_2} | -0,03 | -1,18 | 8,44 | -2,27 |
| β_{7_3} | -2,02 | -1,97 | 2,5 | -7,61 |

Dari Tabel 4.5 dapat diperoleh informasi mengenai efek parameter terhadap pengeluaran rumah tangga di wilayah perkotaan. Variabel yang memiliki pengaruh yang semakin besar pada pengeluaran rumah tangga untuk konsumsi rokok yang semakin tinggi adalah variabel pendapatan (X_1), tingkat pendidikan kepala rumah tangga menengah-rendah (X_{2_1}), jumlah anggota rumah tangga (X_3), umur (X_5), dan jenis kelamin kepala rumah tangga (X_6). Hal ini bisa dilihat dari nilai estimator yang memiliki kecenderungan semakin tinggi seiring dengan kenaikan kuantil. Sedangkan yang memberikan kecenderungan nilai estimator yang semakin kecil di setiap kenaikan kuantil adalah tingkat pendidikan kepala rumah tangga tinggi-rendah (X_{2_2}), rata-rata pengeluaran per kapita (X_4), sektor pekerjaan di sektor sekunder-primer (X_{7_1}), sektor pekerjaan di sektor tersier-primer (X_{7_2}), dan sektor pekerjaan di sektor tidak bekerja-primer (X_{7_3}). Kecenderungan estimator yang semakin kecil menunjukkan bahwa pada kenaikan pengeluaran rumah tangga untuk konsumsi rokok maka pengaruh dari variabel tersebut akan semakin kecil.

Tabel 4.6 Estimasi Parameter Regresi Kuantil Tersensor di Wilayah Pedesaan

| Estimasi Parameter | Estimator Kuantil | | | |
|-----------------------|----------------------|--------|-------|-------|
| | 0,25 | 0,5 | 0,75 | 0,9 |
| β_0 | -9,63 | -5,91 | 7,06 | 11,45 |
| β_1 | 0,0006 | 0,007 | 0,013 | 0,02 |
| β_{2_1} | -1,54 | -12,59 | -9,24 | 8,75 |
| β_{2_2} | -0,52 | 0,18 | 1,29 | 0,19 |
| β_3 | 1,75 | 4,99 | 5,88 | 9,61 |
| β_4 | 0,17 | 0,46 | 0,37 | 0,53 |
| β_5 | -0,04 | -0,34 | -0,22 | -0,36 |
| β_6 | -0,55 | -14,98 | -7,04 | -0,53 |
| β_{7_1} | -9,63 | -5,91 | 7,06 | 11,45 |
| β_{7_2} | 0,0006 | 0,007 | 0,013 | 0,02 |
| β_{7_3} | -1,54 | -12,59 | -9,24 | 8,75 |

Dari Tabel 4.6 dapat diperoleh informasi mengenai efek parameter terhadap pengeluaran rumah tangga di wilayah pedesaan. Variabel yang memiliki pengaruh yang semakin besar pada pengeluaran rumah tangga untuk konsumsi rokok yang semakin tinggi adalah variabel pendapatan (X_1), tingkat pendidikan kepala rumah tangga tinggi-rendah (X_{2_2}), jumlah anggota rumah tangga (X_3), rata-rata pengeluaran per kapita (X_4), sektor pekerjaan di sektor sekunder-primer (X_{7_1}), dan sektor pekerjaan di sektor tersier-primer (X_{7_2}). Hal ini bisa dilihat dari nilai estimator yang memiliki kecenderungan semakin tinggi seiring dengan kenaikan kuantil. Sedangkan variabel yang memberikan kecenderungan nilai estimator yang semakin kecil di setiap kenaikan kuantil adalah tingkat pendidikan kepala rumah tangga menengah-rendah (X_{2_1}), umur (X_5), jenis kelamin kepala rumah tangga (X_6), dan sektor pekerjaan di sektor tidak bekerja-primer (X_{7_3}). Kecenderungan estimator yang semakin kecil menunjukkan bahwa pada kenaikan pengeluaran rumah tangga untuk konsumsi rokok maka pengaruh dari variabel tersebut akan semakin kecil.

Tabel 4.7 Estimasi Parameter Regresi Kuantil di Wilayah Gabungan

| Estimasi Parameter | Estimator Kuantil | | | | |
|-----------------------|----------------------|---------|--------|--------|--------|
| | 0,1 | 0,25 | 0,5 | 0,75 | 0,9 |
| β_0 | 0,00 | -0,65 | 13,60 | 23,51 | 32,54 |
| β_1 | 0,00 | 0,0001 | 0,004 | 0,01 | 0,014 |
| β_{2_1} | 0,00 | -1,14 | -14,85 | -14,65 | -8,29 |
| β_{2_2} | 0,00 | -0,46 | -1,19 | -3,27 | -2,56 |
| β_3 | 0,00 | 0,70 | 4,52 | 6,44 | 10,74 |
| β_4 | 0,00 | 0,00001 | 0,003 | 0,008 | 0,017 |
| β_5 | 0,00 | -0,02 | -0,32 | -0,27 | -0,35 |
| β_6 | 0,00 | -0,87 | -12,22 | -17,68 | -7,05 |
| β_{7_1} | 0,00 | -0,21 | 2,63 | 1,60 | -5,48 |
| β_{7_2} | 0,00 | -0,41 | -1,04 | 0,94 | -0,10 |
| β_{7_3} | 0,00 | -0,46 | -3,02 | -3,83 | -7,83 |
| β_8 | 0,00 | -0,37 | -3,12 | -6,01 | -10,54 |

Dari Tabel 4.7 dapat diperoleh informasi mengenai efek parameter terhadap pengeluaran rumah tangga di wilayah gabungan. Berbeda dengan hasil analisis metode regresi kuantil tersensor, pada kuantil bawah (0,1) metode regresi kuantil tidak menghasilkan estimasi parameter seperti halnya metode regresi kuantil tersensor. Melihat dari nilai estimator yang dihasilkan maka terdapat variabel yang memiliki kecenderungan yang semakin besar atau semakin kecil di setiap kenaikan kuantil, namun terdapat juga variabel yang tidak memiliki kecenderungan yang tetap di setiap kenaikan kuantil. Variabel yang memiliki kecenderungan semakin besar di setiap kenaikan kuantil adalah variabel pendapatan (X_1), jumlah anggota rumah tangga (X_3), dan rata-rata pengeluaran per kapita (X_4). Variabel yang memiliki kecenderungan semakin kecil di setiap kenaikan kuantil adalah sektor pekerjaan di sektor tidak bekerja-primer (X_{7_3}) dan wilayah tempat tinggal (X_8). Sedangkan variabel yang tidak memiliki kecenderungan semakin besar atau semakin kecil dalam setiap kenaikan kuantil adalah tingkat pendidikan kepala rumah tangga menengah-rendah (X_{2_1}), tingkat pendidikan kepala rumah tangga tinggi-rendah (X_{2_2}), umur (X_5), jenis kelamin kepala rumah tangga (X_6), sektor pekerjaan di sektor sekunder-primer (X_{7_1}), dan sektor pekerjaan di sektor tersier-primer (X_{7_2}).

Tabel 4.8 Estimasi Parameter Regresi Kuantil di Wilayah Perkotaan

| Estimasi Parameter | Estimator Kuantil | | | |
|-----------------------|----------------------|---------|---------|---------|
| | 0,25 | 0,5 | 0,75 | 0,9 |
| β_0 | -4,505 | -4,715 | 5,260 | 14,624 |
| β_1 | 0,001 | 0,006 | 0,011 | 0,020 |
| β_{2_1} | -4,728 | -14,349 | -19,289 | -0,839 |
| β_{2_2} | -2,244 | -4,507 | -8,063 | -5,507 |
| β_3 | 1,669 | 3,406 | 5,278 | 8,606 |
| β_4 | 0,270 | 0,600 | 0,488 | 0,681 |
| β_5 | -0,084 | -0,039 | 0,032 | 0,113 |
| β_6 | -6,465 | -21,239 | -24,022 | -8,897 |
| β_{7_1} | -0,011 | 3,020 | 2,773 | -20,719 |
| β_{7_2} | -0,755 | -0,153 | 9,331 | -5,361 |
| β_{7_3} | -0,764 | -0,930 | 2,585 | -7,341 |

Dari Tabel 4.8 dapat diperoleh informasi mengenai efek parameter terhadap pengeluaran rumah tangga di wilayah perkotaan. Melihat dari nilai estimator yang dihasilkan maka terdapat variabel yang memiliki kecenderungan yang semakin besar di setiap kenaikan kuantil, namun terdapat juga variabel yang tidak memiliki kecenderungan yang tetap di setiap kenaikan kuantil. Variabel yang memiliki kecenderungan semakin besar di setiap kenaikan kuantil adalah variabel pendapatan (X_1), jumlah anggota rumah tangga (X_3), dan rata-rata pengeluaran per kapita (X_4), dan umur (X_5). Sedangkan variabel yang tidak memiliki kecenderungan semakin besar atau semakin kecil dalam setiap kenaikan kuantil adalah tingkat pendidikan kepala rumah tangga menengah-rendah (X_{2_1}), tingkat pendidikan kepala rumah tangga tinggi-rendah (X_{2_2}), jenis kelamin kepala rumah tangga (X_6), sektor pekerjaan di sektor sekunder-primer (X_{7_1}), sektor pekerjaan di sektor tersier-primer (X_{7_2}), dan sektor pekerjaan di sektor tidak bekerja-primer (X_{7_3}).

Tabel 4.9 Estimasi Parameter Regresi Kuantil di Wilayah Pedesaan

| Estimasi Parameter | Estimator Kuantil | | | |
|-----------------------|----------------------|--------|--------|-------|
| | 0,25 | 0,5 | 0,75 | 0,9 |
| β_0 | -2,79 | -4,40 | 7,73 | 13,77 |
| β_1 | 0,0003 | 0,007 | 0,01 | 0,02 |
| β_{2_1} | -0,92 | -10,87 | -11,49 | 8,67 |
| β_{2_2} | -0,06 | 1,20 | 0,93 | 0,18 |
| β_3 | 0,78 | 4,01 | 6,36 | 9,40 |
| β_4 | 0,06 | 0,43 | 0,37 | 0,51 |
| β_5 | -0,02 | -0,23 | -0,30 | -0,37 |
| β_6 | -0,59 | -7,61 | -7,03 | 0,38 |
| β_{7_1} | 0,12 | 1,97 | 2,40 | 1,99 |
| β_{7_2} | -0,53 | 1,80 | 3,67 | 5,11 |
| β_{7_3} | -0,50 | -4,06 | -4,66 | -0,90 |

Dari Tabel 4.9 dapat diperoleh informasi mengenai efek parameter terhadap pengeluaran rumah tangga di wilayah perkotaan. Melihat dari nilai estimator yang dihasilkan maka terdapat variabel yang memiliki kecenderungan yang semakin besar di setiap kenaikan kuantil, namun terdapat juga variabel yang tidak memiliki kecenderungan yang tetap di setiap kenaikan kuantil. Variabel yang memiliki kecenderungan semakin besar di setiap kenaikan kuantil adalah variabel pendapatan (X_1), jumlah anggota rumah tangga (X_3), dan rata-rata pengeluaran per kapita (X_4), dan sektor pekerjaan di sektor tersier-primer (X_{7_2}). Variabel yang memiliki kecenderungan semakin kecil di setiap kenaikan kuantil adalah umur (X_5). Sedangkan variabel yang tidak memiliki kecenderungan semakin besar atau semakin kecil dalam setiap kenaikan kuantil adalah tingkat pendidikan kepala rumah tangga menengah-rendah (X_{2_1}), tingkat pendidikan kepala rumah tangga tinggi-rendah (X_{2_2}), jenis kelamin kepala rumah tangga (X_6), sektor pekerjaan di sektor sekunder-primer (X_{7_1}), dan sektor pekerjaan di sektor tidak bekerja-primer (X_{7_3}).

Selanjutnya akan dilakukan penghitungan nilai RMSE untuk menguji kebaikan dari model yang telah terbentuk. Sebagai bahan perbandingan maka metode regresi kuantil tersensor akan dibandingkan dengan metode regresi kuantil. Tabel nilai RMSE kedua metode tersebut akan disajikan pada Tabel 4.10.

Tabel 4.10 RMSE Regresi Kuantil Tersensor dan Regresi Kuantil

| Kuantil | Regresi Kuantil Tersensor | | | Regresi Kuantil | | |
|---------|---------------------------|-----------|----------|-----------------|-----------|----------|
| | Gabungan | Perkotaan | Pedesaan | Gabungan | Perkotaan | Pedesaan |
| 0,1 | 49,9 | - | - | - | - | - |
| 0,25 | 7,38 | 17,24 | 9,10 | 1,47 | 5,79 | 8,96 |
| 0,5 | 24,77 | 28,65 | 26,48 | 26,19 | 29,46 | 26,72 |
| 0,75 | 55,45 | 61,66 | 58,29 | 55,59 | 62,53 | 59,34 |
| 0,9 | 88,09 | 107,62 | 90,69 | 92,60 | 108,30 | 94,74 |

Pada Tabel 4.10 dapat diperoleh informasi mengenai kebaikan model pengeluaran rumah tangga untuk konsumsi rokok di wilayah gabungan (perkotaan dan pedesaan), wilayah perkotaan, dan wilayah pedesaan menggunakan metode regresi kuantil tersensor yang akan dibandingkan dengan regresi kuantil. Perbandingan nilai RMSE pada ketiga model menggunakan metode regresi kuantil tersensor dan metode regresi kuantil memberikan hasil bahwa pada kuantil 0,25 nilai RMSE terkecil dimiliki oleh metode regresi kuantil, sedangkan pada kuantil 0,5, 0,75, dan 0,9 nilai RMSE terkecil dimiliki oleh metode regresi kuantil tersensor. Sehingga dapat dikatakan bahwa metode regresi kuantil tersensor memiliki kebaikan dalam pemodelan pengeluaran rumah tangga untuk konsumsi rokok pada kuantil 0,5, 0,75, dan 0,9. Namun demikian pada model pengeluaran rumah tangga untuk konsumsi rokok di wilayah gabungan khususnya pada kuantil bawah (0,1) metode regresi kuantil tersensor dapat menghasilkan estimator yang tidak dapat dihasilkan oleh metode regresi kuantil.

Berdasarkan hasil analisis menggunakan metode regresi kuantil tersensor yang dibandingkan dengan metode regresi kuantil diperoleh hasil bahwa pada model pengeluaran rumah tangga untuk konsumsi rokok di wilayah gabungan dengan menggunakan metode regresi kuantil tersensor menghasilkan estimasi parameter pada kuantil bawah (0,1), sedangkan metode regresi kuantil tidak dapat menghasilkan estimasi parameter pada kuantil tersebut. Hal ini bisa diartikan bahwa metode regresi kuantil tersensor dapat memodelkan pengeluaran rumah tangga untuk konsumsi rokok pada kuantil 0,1. Sedangkan jika melihat dari hasil estimasi secara keseluruhan baik pada model di wilayah gabungan, perkotaan, dan pedesaan, kedua metode tersebut dapat digunakan untuk mengetahui kecenderungan variabel prediktor terhadap pengeluaran rumah tangga untuk konsumsi rokok. Namun demikian, estimasi menggunakan metode regresi kuantil

menghasilkan estimator pada beberapa variabel yang tidak mengalami kecenderungan semakin besar atau semakin kecil pada setiap kenaikan kuantil. Hal ini berbeda pada estimator yang dihasilkan menggunakan metode regresi kuantil tersensor yang menghasilkan estimator yang memiliki kecenderungan semakin besar atau semakin kecil pada setiap kenaikan kuantil sehingga lebih mudah diketahui hubungan yang terbentuk diantara pengeluaran rumah tangga untuk konsumsi rokok dan variabel prediktor yang digunakan. Selain itu, melihat dari nilai RMSE yang diperoleh diantara metode regresi kuantil tersensor dan regresi kuantil yang diperoleh hasil bahwa metode regresi kuantil tersensor memiliki kebaikan dalam pemodelan pengeluaran rumah tangga untuk konsumsi rokok pada kuantil 0,5, 0,75, dan 0,9. Sehingga dapat diperoleh kesimpulan bahwa metode regresi kuantil tersensor lebih tepat digunakan untuk memodelkan pengeluaran rumah tangga untuk konsumsi rokok dibandingkan dengan metode regresi kuantil biasa.

4.3.3 Model Regresi Kuantil Tersensor

Berikut akan disajikan model regresi kuantil tersensor yang terbentuk di wilayah gabungan (perkotaan dan pedesaan), wilayah perkotaan, dan wilayah pedesaan. Model regresi kuantil tersensor di wilayah gabungan (perkotaan dan pedesaan) adalah sebagai berikut.

$$\hat{y}_{(0,1)} = \begin{cases} -80,02 + 0,0017X_1 - 32,32X_{21} - 12,4X_{22} + 6,76X_3 - 0,0049X_4 & y^* > 0 \\ +0,11X_5 - 14,1X_6 + 27,2X_{71} + 1,63X_{72} + 28,62X_{73} + 2,22X_8 & \\ 0 & y^* \leq 0 \end{cases}$$

$$\hat{y}_{(0,25)} = \begin{cases} -4,51 + 0,001X_1 - 12,33X_{21} - 1,67X_{22} + 2,27X_3 - 0,0009X_4 & y^* > 0 \\ -0,06X_5 - 20,48X_6 + 0,82X_{71} - 0,33X_{72} + 0,01X_{73} - 1,99X_8 & \\ 0 & y^* \leq 0 \end{cases}$$

$$\hat{y}_{(0,5)} = \begin{cases} 17,61 + 0,005X_1 - 27,58X_{21} - 2,02X_{22} + 5,09X_3 + 0,004X_4 & y^* > 0 \\ -0,48X_5 - 15,77X_6 + 1,22X_{71} - 2,009X_{72} - 3,85X_{73} - 4,55X_8 & \\ 0 & y^* \leq 0 \end{cases}$$

$$\hat{y}_{(0,75)} = \begin{cases} 23,50 + 0,01X_1 - 16,89X_{21} - 2,39X_{22} + 6,1X_3 + 0,009X_4 & y^* > 0 \\ -0,24X_5 - 16,96X_6 + 1,8X_{71} + 0,94X_{72} - 3,35X_{73} - 5,65X_8 & \\ 0 & y^* \leq 0 \end{cases}$$

$$\hat{y}_{(0,9)} = \begin{cases} 33,1 + 0,01X_1 - 8,81X_{21} - 1,49X_{22} + 10,73X_3 + 0,02X_4 & y^* > 0 \\ -0,37X_5 + 4,65X_6 - 6,9X_{71} - 0,08X_{72} - 3,35X_{73} - 9,89X_8 & \\ 0 & y^* \leq 0 \end{cases}$$

Berdasarkan model regresi kuantil tersensor di wilayah gabungan dapat diperoleh informasi bahwa pada kuantil 0,9 terdapat koefisien yang bertanda positif yaitu pendapatan, jumlah anggota rumah tangga, dan rata-rata pengeluaran per kapita, artinya variabel tersebut memberikan pengaruh secara positif terhadap pengeluaran rumah tangga untuk konsumsi rokok. Dengan kata lain jika terdapat perubahan pendapatan sebesar satu satuan maka akan meningkatkan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,01 rupiah, jika terdapat perubahan jumlah anggota rumah tangga sebesar satu satuan maka akan meningkatkan pengeluaran rumah tangga untuk konsumsi rokok sebesar 10,73 rupiah, dan jika terdapat perubahan rata-rata pengeluaran per kapita sebesar satu satuan maka akan meningkatkan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,02 rupiah. Sedangkan koefisien yang bertanda negatif yaitu umur kepala rumah tangga artinya jika terdapat perubahan umur kepala rumah tangga sebesar satu satuan maka akan menurunkan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,37 rupiah. Selanjutnya akan dilihat koefisien variabel yang berbentuk kategorik dengan melihat nilai odds ratio yang dihasilkan yaitu dengan menghitung nilai eksponen dari masing-masing koefisien. Pada variabel tingkat pendidikan dapat dilihat bahwa pada jenjang pendidikan SLTP dan SMU memiliki kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,00015 kali dibandingkan dengan jenjang pendidikan SD atau tidak tamat SD. Sedangkan pada jenjang pendidikan Perguruan Tinggi memiliki kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,22 kali dibandingkan dengan jenjang pendidikan SD atau tidak tamat SD. Pada variabel jenis kelamin kepala rumah tangga diperoleh informasi bahwa rumah tangga yang memiliki kepala rumah tangga laki-laki mengalami kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar 104,6 kali dibandingkan dengan rumah tangga yang memiliki kepala rumah tangga perempuan. Pada variabel sektor pekerjaan dapat diperoleh

informasi bahwa pada sektor pekerjaan sekunder memiliki kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,001 kali dibandingkan dengan sektor pekerjaan primer, pada sektor pekerjaan tersier memiliki kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,92 kali dibandingkan dengan sektor pekerjaan primer, sedangkan pada sektor pekerjaan kategori tidak bekerja memiliki kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,03 kali dibandingkan dengan sektor pekerjaan primer. Pada variabel wilayah dapat diperoleh informasi bahwa wilayah perkotaan memiliki kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar 5,08 kali dibandingkan dengan wilayah pedesaan. Dengan cara yang sama dapat dilihat interpretasi model pada kuantil lainnya.

Model regresi kuantil tersensor di wilayah perkotaan adalah sebagai berikut.

$$\begin{aligned}\hat{y}_{(0,25)} &= \begin{cases} -36,8 + 0,0041X_1 - 21,91X_{21} - 3,96X_{22} + 4,41X_3 + 0,82X_4 & y^* > 0 \\ -0,12X_5 - 30,22X_6 - 1,28X_{71} - 0,03X_{72} - 2,02X_{73} & \\ 0 & y^* \leq 0 \end{cases} \\ \hat{y}_{(0,5)} &= \begin{cases} -3,34 + 0,005X_1 - 14,36X_{21} - 4,24X_{22} + 3,44X_3 + 0,59X_4 & y^* > 0 \\ -0,05X_5 - 22,02X_6 + 2,58X_{71} - 1,18X_{72} - 1,97X_{73} & \\ 0 & y^* \leq 0 \end{cases} \\ \hat{y}_{(0,75)} &= \begin{cases} 6,78 + 0,01X_1 - 21,27X_{21} - 7,61X_{22} + 5,1X_3 + 0,46X_4 & y^* > 0 \\ +0,05X_5 - 23,94X_6 + 2,44X_{71} + 8,44X_{72} + 2,5X_{73} & \\ 0 & y^* \leq 0 \end{cases} \\ \hat{y}_{(0,9)} &= \begin{cases} 14,93 + 0,02X_1 - 3,74X_{21} - 8,25X_{22} + 8,54X_3 + 0,67X_4 & y^* > 0 \\ +0,1X_5 - 14,18X_6 - 15,59X_{71} - 2,27X_{72} - 7,61X_{73} & \\ 0 & y^* \leq 0 \end{cases}\end{aligned}$$

Berdasarkan model regresi kuantil tersensor di wilayah Perkotaan dapat diperoleh informasi bahwa pada kuantil 0,9 terdapat koefisien yang bertanda positif yaitu pendapatan, jumlah anggota rumah tangga, rata-rata pengeluaran per kapita, dan umur kepala rumah tangga, artinya variabel tersebut memberikan pengaruh secara positif terhadap pengeluaran rumah tangga untuk konsumsi rokok. Dengan kata lain jika terdapat perubahan pendapatan sebesar satu satuan

maka akan meningkatkan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,02 rupiah, jika terdapat perubahan jumlah anggota rumah tangga sebesar satu satuan maka akan meningkatkan pengeluaran rumah tangga untuk konsumsi rokok sebesar 8,54 rupiah, jika terdapat perubahan rata-rata pengeluaran per kapita sebesar satu satuan maka akan meningkatkan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,67 rupiah, dan jika terdapat perubahan umur kepala rumah tangga sebesar satu satuan maka akan meningkatkan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,1 rupiah. Selanjutnya akan dilihat koefisien variabel yang berbentuk kategorik dengan melihat nilai odds ratio yang dihasilkan yaitu dengan menghitung nilai eksponen dari masing-masing koefisien. Pada variabel tingkat pendidikan dapat dilihat bahwa pada jenjang pendidikan SLTP dan SMU memiliki kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,02 kali dibandingkan dengan jenjang pendidikan SD atau tidak tamat SD. Sedangkan pada jenjang pendidikan Perguruan Tinggi memiliki kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,0003 kali dibandingkan dengan jenjang pendidikan SD atau tidak tamat SD. Pada variabel jenis kelamin kepala rumah tangga diperoleh informasi bahwa rumah tangga yang memiliki kepala rumah tangga laki-laki mengalami kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar $6,9 \times 10^{-07}$ kali dibandingkan dengan rumah tangga yang memiliki kepala rumah tangga perempuan. Pada variabel sektor pekerjaan dapat diperoleh informasi bahwa pada sektor pekerjaan sekunder memiliki kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar $1,7 \times 10^{-07}$ kali dibandingkan dengan sektor pekerjaan primer, pada sektor pekerjaan tersier memiliki kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,1 kali dibandingkan dengan sektor pekerjaan primer, sedangkan pada sektor pekerjaan kategori tidak bekerja memiliki kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,0005 kali dibandingkan dengan sektor pekerjaan primer. Dengan cara yang sama dapat dilihat interpretasi model pada kuantil lainnya.

Model regresi kuantil tersensor di wilayah pedesaan adalah sebagai berikut.

$$\begin{aligned}\hat{y}_{(0,25)} &= \begin{cases} -9,63 + 0,0006X_1 - 1,54X_{21} - 0,52X_{22} + 1,75X_3 + 0,17X_4 & y^* > 0 \\ -0,04X_5 - 0,55X_6 + 1,4X_{71} + 0,48X_{72} + 0,81X_{73} & y^* \leq 0 \\ 0 & y^* \leq 0 \end{cases} \\ \hat{y}_{(0,5)} &= \begin{cases} -5,91 + 0,007X_1 - 12,59X_{21} + 0,18X_{22} + 4,99X_3 + 0,46X_4 & y^* > 0 \\ -0,34X_5 - 14,98X_6 + 4,18X_{71} + 1,69X_{72} - 3,59X_{73} & y^* \leq 0 \\ 0 & y^* \leq 0 \end{cases} \\ \hat{y}_{(0,75)} &= \begin{cases} 7,06 + 0,013X_1 - 9,24X_{21} + 1,29X_{22} + 5,88X_3 + 0,37X_4 & y^* > 0 \\ -0,22X_5 - 7,04X_6 + 5,38X_{71} + 3,98X_{72} - 4,59X_{73} & y^* \leq 0 \\ 0 & y^* \leq 0 \end{cases} \\ \hat{y}_{(0,9)} &= \begin{cases} 11,45 + 0,02X_1 + 8,75X_{21} + 0,19X_{22} + 9,61X_3 + 0,53X_4 & y^* > 0 \\ -0,36X_5 - 0,53X_6 + 10,87X_{71} + 4,88X_{72} - 0,66X_{73} & y^* \leq 0 \\ 0 & y^* \leq 0 \end{cases}\end{aligned}$$

Berdasarkan model regresi kuantil tersensor di wilayah Pedesaan dapat diperoleh informasi bahwa pada kuantil 0,9 terdapat koefisien yang bertanda positif yaitu pendapatan, jumlah anggota rumah tangga, dan rata-rata pengeluaran per kapita, artinya jika terdapat perubahan pendapatan sebesar satu satuan maka akan meningkatkan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,02 rupiah, jika terdapat perubahan jumlah anggota rumah tangga sebesar satu satuan maka akan meningkatkan pengeluaran rumah tangga untuk konsumsi rokok sebesar 9,61 rupiah, dan jika terdapat perubahan rata-rata pengeluaran per kapita sebesar satu satuan maka akan meningkatkan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,53 rupiah. Sedangkan koefisien yang bertanda negatif yaitu umur kepala rumah tangga artinya jika terdapat perubahan umur kepala rumah tangga sebesar satu satuan maka akan menurunkan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,36 rupiah. Selanjutnya akan dilihat koefisien variabel yang berbentuk kategorik dengan melihat nilai odds ratio yang dihasilkan yaitu dengan menghitung nilai eksponen dari masing-masing koefisien. Pada variabel tingkat pendidikan dapat dilihat bahwa pada jenjang pendidikan SLTP dan SMU memiliki kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar 6.310 kali dibandingkan dengan jenjang

pendidikan SD atau tidak tamat SD. Sedangkan pada jenjang pendidikan Perguruan Tinggi memiliki kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar 1,2 kali dibandingkan dengan jenjang pendidikan SD atau tidak tamat SD. Pada variabel jenis kelamin kepala rumah tangga diperoleh informasi bahwa rumah tangga yang memiliki kepala rumah tangga laki-laki mengalami kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,6 kali dibandingkan dengan rumah tangga yang memiliki kepala rumah tangga perempuan. Pada variabel sektor pekerjaan dapat diperoleh informasi bahwa pada sektor pekerjaan sekunder memiliki kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar 52.575 kali dibandingkan dengan sektor pekerjaan primer, pada sektor pekerjaan tersier memiliki kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar 131 kali dibandingkan dengan sektor pekerjaan primer, sedangkan pada sektor pekerjaan kategori tidak bekerja memiliki kecenderungan melakukan pengeluaran rumah tangga untuk konsumsi rokok sebesar 0,5 kali dibandingkan dengan sektor pekerjaan primer. Dengan cara yang sama dapat dilihat interpretasi model pada kuantil lainnya.

BAB 5

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis yang telah dilakukan maka dapat diperoleh kesimpulan sebagai berikut.

1. Untuk mendapatkan estimator $\beta_0(\theta)$ regresi kuantil tersensor menggunakan *loss function* dengan meminimalkan persamaan dibawah ini menggunakan algoritma optimasi.

$$Q_n(\beta; \theta) = \frac{1}{n} \sum_{i=1}^n \rho_{\theta}(y_i - \max\{0, \mathbf{x}_i^T \beta\})$$

Proses optimasi yang digunakan adalah menggunakan tiga tahap algoritma yaitu mengambil sub sampel dengan cara pemisahan dari probabilitas tersensor, dan melakukan dua kali estimasi menggunakan regresi kuantil.

2. Berdasarkan hasil simulasi yang dilakukan untuk melihat performa estimator regresi kuantil tersensor diperoleh hasil bahwa pada kondisi terdapat variabel tersensor maka regresi kuantil tersensor memiliki performa yang lebih baik dibandingkan dengan regresi kuantil. Hal ini dapat dilihat dari nilai RMSE intersep masing-masing metode yang menunjukkan bahwa metode regresi kuantil tersensor memiliki nilai RMSE yang lebih kecil dibandingkan dengan metode regresi kuantil.
3. Berdasarkan hasil pemodelan regresi kuantil tersensor dapat diperoleh kesimpulan bahwa:
 - a. Pada model gabungan variabel yang memberikan pengaruh yang besar untuk konsumsi rokok yang tinggi adalah variabel pendapatan (X_1), tingkat pendidikan kepala rumah tangga menengah-rendah (X_{2_1}), tingkat pendidikan kepala rumah tangga tinggi-rendah (X_{2_2}), jumlah anggota rumah tangga (X_3), rata-rata pengeluaran per kapita (X_4), dan jenis kelamin kepala rumah tangga (X_6). Sedangkan variabel yang memberikan pengaruh yang kecil pada kenaikan pengeluaran konsumsi rokok adalah variabel umur (X_5), sektor pekerjaan di sektor

sekunder-primer (X_{7_1}), sektor pekerjaan di sektor tersier-primer (X_{7_2}), sektor pekerjaan di sektor tidak bekerja-primer (X_{7_3}), dan wilayah tempat tinggal (X_8).

- b. Pada model di wilayah Perkotaan variabel yang memiliki pengaruh yang semakin besar pada pengeluaran rumah tangga untuk konsumsi rokok yang semakin tinggi adalah variabel pendapatan (X_1), tingkat pendidikan kepala rumah tangga menengah-rendah (X_{2_1}), jumlah anggota rumah tangga (X_3), umur (X_5), dan jenis kelamin kepala rumah tangga (X_6). Sedangkan yang memberikan kecenderungan nilai estimator yang semakin kecil di setiap kenaikan kuantil adalah tingkat pendidikan kepala rumah tangga tinggi-rendah (X_{2_2}), rata-rata pengeluaran per kapita (X_4), sektor pekerjaan di sektor sekunder-primer (X_{7_1}), sektor pekerjaan di sektor tersier-primer (X_{7_2}), dan sektor pekerjaan di sektor tidak bekerja-primer (X_{7_3}).
- c. Pada model di wilayah Pedesaan variabel yang memiliki pengaruh yang semakin besar pada pengeluaran rumah tangga untuk konsumsi rokok yang semakin tinggi adalah variabel pendapatan (X_1), tingkat pendidikan kepala rumah tangga tinggi-rendah (X_{2_2}), jumlah anggota rumah tangga (X_3), rata-rata pengeluaran per kapita (X_4), sektor pekerjaan di sektor sekunder-primer (X_{7_1}), dan sektor pekerjaan di sektor tersier-primer (X_{7_2}). Sedangkan variabel yang memberikan kecenderungan nilai estimator yang semakin kecil di setiap kenaikan kuantil adalah tingkat pendidikan kepala rumah tangga menengah-rendah (X_{2_1}), umur (X_5), jenis kelamin kepala rumah tangga (X_6), dan sektor pekerjaan di sektor tidak bekerja-primer (X_{7_3}).
- d. Pada hasil estimasi menggunakan metode regresi kuantil menghasilkan estimator pada beberapa variabel yang tidak mengalami kecenderungan semakin besar atau semakin kecil pada setiap kenaikan kuantil. Hal ini berbeda pada estimator yang dihasilkan menggunakan metode regresi kuantil tersensor yang menghasilkan estimator yang memiliki kecenderungan semakin besar atau semakin kecil pada setiap

kenaikan kuantil sehingga lebih mudah diketahui hubungan yang terbentuk diantara pengeluaran rumah tangga untuk konsumsi rokok dan variabel prediktor yang digunakan. Selain itu, melihat dari nilai RMSE yang diperoleh diantara kedua metode diperoleh hasil bahwa metode regresi kuantil tersensor memiliki kebaikan dalam pemodelan pengeluaran rumah tangga untuk konsumsi rokok pada kuantil 0,5, 0,75, dan 0,9. Hal ini dapat dilihat dari nilai RMSE metode regresi kuantil tersensor yang lebih kecil pada kuantil tersebut.

5.2 Saran

Saran-saran yang dapat disampaikan berdasarkan hasil penelitian yang ada adalah sebagai berikut.

1. Perlu dikaji lebih dalam terhadap estimasi parameter secara komputasi terutama tentang tahapan algoritma yang digunakan untuk melakukan optimasi dalam estimasi parameter.
2. Dari hasil simulasi dan analisis yang telah dilakukan, ditemukan bahwa estimator regresi kuantil tersensor cukup baik untuk mengestimasi model pada jumlah sampel yang berbeda dan distribusi *error* adalah normal, namun belum diketahui bagaimana performa pada kondisi yang lainnya. Oleh karena itu disarankan untuk melakukan simulasi dengan kemungkinan lainnya.

(halaman ini sengaja dikosongkan)

DAFTAR PUSTAKA

- Agresti, A. (2002), *Categorical Data Analysis*, New York: John Wiley & Sons.
- Babolian, H., R., Karim, M., S. (2010), Factors Affecting Milk Consumption Among School Children in Urban and Rural Areas of Selangor Malaysia, *International Food Research Journal* 17:651-660.
- Badan Pusat Statistik [BPS], (2015), *Survey Sosial Ekonomi Nasional*, Publikasi Badan Pusat Statistik, Jakarta.
- Buchinsky, M. (1994), "Changes in US Wage Structure 1963-87: An Application of Quantile Regression", *Econometrica*, 62, 405-458.
- Buhai, I. S. (2005), "Quantile Regression: Overview and Selected Applications". *Ad Astra*, Vol. 4, 1-17.
- Cahyaningsih, A. (2011), *Pendekatan Tobit Model dan Double Hurdle Dalam Pemodelan Pengeluaran Konsumsi Rokok di Kalimantan Timur*, Tesis Master, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia.
- Chai, T., dan Draxler, R. R. (2014). "Root Mean Square Error (RMSE) or Mean Absolute Error (MAE): Arguments Against Avoiding RMSE in The Literature". *Geoscientific Model Development*, Vol. 7, 1247-1250.
- Chen, C. dan Wei, Y. (2005), Computational Issues for Quantile Regression, *The Indian Journal of Statistics*, Volume 67, Part 2, pp 399-417.
- Chen, C. (2005), An Introduction to Quantile Regression and The QUANTREG Procedure. Retrieved October 20th, 2014, from <http://www2.sas.com/proceedingd/sugi30/213-30.pdf>
- Chernozhukov, V. dan Hong, H. (2002), Three-Step Censored Quantile Regression and Extramarital Affairs, *Journal of the American Statistical Association*.
- Davino, C., Furno, M., dan Vistocco, D. (2014), *Quantile Regression: Theory and Application*, John Wiley and Sons, Ltd, UK
- Dinas Kesehatan Kabupaten Kudus, (2015), *Inilah 4 Bahaya Merokok Bagi Kesehatan Tubuh*, Kudus, Indonesia

- Faidah, D., Y. (2012), *Model Tobit Spasial Pada Faktor-Faktor yang Mempengaruhi Tingkat Pengangguran Terbuka (tpt) Perempuan*, Tesis Master, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia.
- Fitzenberger, B. (1997), *Computational Aspects of Censored Quantile Regression, IMS Lecturer Note- Monograph Series (1997) Volume 31.*
- Friederichs, P dan Hense, A. (2006), *A Statistical Downscaling of Extreme Precipitation Events Using Quantile Regression*, Meteorological Institute, University of Bonn, Bonn, Germany.
- Furno, M. (2007), *Parameter Instability in Quantile Regression*, Statistical Modelling, 7(4):345-362.
- Global Adult Tobacco Survey [GATS], (2011), *Global Adult Tobacco Survey: Indonesia Report 2011*, World Health Organization, Regional Office for South East Asia.
- Greene, W.H. (2008). *Econometrics Analysis*, 6th edition. Prentice Hall, New Jersey.
- Gujarati, (2004), *Basic Econometrics :Fourth Edition*. New York :McGraw Hill.
- Gustavsen, G.,W., Jolliffe, D., dan Rickertsen, K. (2008), *Censored Quantile Regression and Purchases of Ice Cream, Selected Paper Prepared for Presentation at The American Agricultural Economics Association Annual Meeting, Orlando, Florida, July 27-29, 2008.*
- Gustavsen, G.,W. dan Rickertsen, K. (2013), *Adjusting VAT Rates to Promote Healthier Diets in Norway: A Censored Quantile Regression Approach*, *Food Policy* 42(2013) 88-95.
- Hao, L., dan Naiman, D., Q. (2007), *Quantile Regression*, Sage Publication Inc, United State of America.
- Helmi, D., P. (2016), *Analisis Faktor-Faktor Yang Mempengaruhi Perilaku Merokok Pada Rumah Tangga Di Kabupaten Lima Puluh Kota Tahun 2013*. Master thesis, Universitas Andalas
- Hocking, R.(1996), *Methods and Application of Linear Models*, John Wiley & Sons, New York.
- Kementrian Kesehatan RI, 2013, *Riset Kesehatan Dasar (Riskesdas)*, Badan Penelitian dan Pengembangan Kesehatan, Jakarta.

- Koenker, R. (2005), *Quantile Regression, Econometric Society Monographs*, Vol. 38, Cambridge University Press, 349 pp.
- Koenker, R., dan Hallock, K. (2001), Quantile Regression, *Journal of Econometric Perspectives*, Vol. 15, 143-156.
- Koenker, R. dan Bassett, G. (1978), *Regression Quantiles, Econometrica*, Vol.46, No.1. (Jan.,1978), pp.33-50.
- Koenker, R. dan d'Orey, V. (1993), *Computing Regression Quantiles*, J. Roy, Statist. Soc. Ser. C (Appl. Statist), Vol 43, pp.410-414.
- Koenker, R. dan Machado, J. A. F. (1999), "Goodness of Fit and Related Inference Processes for Quantile Regression", *Journal of the American Statistical Association*, Vol.94, 1296-1310.
- Leiker, A. (2012), *A Comparison Study on The Estimation in Tobit Regression Models*, Thesis Master of Science, Kansas State University, Kansas.
- Lusiana, E., D. (2015), *Pemodelan Regresi Tobit Kuantil Bayesian Pada Pengeluaran Rumah Tangga Untuk Konsumsi Rokok*, Tesis Master, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia.
- McBee, M. (2010), Modeling Outcomes Floor or Ceiling Effects: An Introduction to the Tobit Model, *Gift Child Quarterly*, 54(4) 314-320.
- Permana, G. (2013), *Analisis regresi Tobit Pada Permasalahan Pengeluaran Konsumsi Rokok Kota Kediri Tahun 2011*, Skripsi, Universitas Brawijaya, Malang.
- Powell, J.,L. (1984), Least Absolute Deviations Estimation For Censored Regression Model, *Journal of Econometrics* 25 (1984) , 303 – 325. North-Holland.
- Powell, J.,L. (1986), Censored Regression Quantiles, *Journal of Econometrics* 32 (1986) 143-155. North-Holland.
- Robinson, C., Tomek, S., dan Schumaker, R. (2013), "Test of Moderation Effects: Difference in Simple Slopes versus the Interaction Term", *Multiple Linear Regression Viewpoints*, Vol.39 (16-24).
- Suhardi, I.Y., & Llewelyn, R., (2001), Penggunaan Model Regresi Tobit untuk

Menganalisa Faktor-Faktor yang Berpengaruh terhadap Kepuasan Konsumen untuk Jasa Pengangkutan Barang, *Jurnal Manajemen & Kewirausahaan*, 3(2), 06-112.

Tobin, J. (1958), Estimation of Relationships for Limited Dependent Variables, *Econometrica*, 26(1), 24-36.

Wooldridge, J.,M. (2002), *Introduction Econometrics:A Modern Approach Second Edition*, USA.

World bank, www.microdata.worldbank.org, diakses pada 6 Agustus 2016.

Yao, Y., dan Lee, Y. (2010), Another Look at Linear Programming for Feature Selection via Methods of Regularization, Ohio: The Ohio State University.

Zain, I. dan Suhartono, 1997, *Model Regresi Tobit dan Aplikasinya*, Laporan Penelitian LPPM ITS, Surabaya.

LAMPIRAN - LAMPIRAN

Lampiran 1 Data Penelitian

| No | Pendapatan (Rp) | Pendidikan | JART | Rata-rata Pengeluaran per Kapita (Rp) | Umur | Jenis Kelamin | Sektor Pekerjaan | Wilayah | Konsumsi Rokok (Rp) |
|------|--------------------|------------|------|--|------|------------------|---------------------|---------|---------------------------|
| 1 | 1000000 | 1 | 7 | 883310 | 53 | 1 | 3 | 1 | 70000 |
| 2 | 800000 | 1 | 4 | 307500 | 37 | 1 | 1 | 1 | 98000 |
| 3 | 700000 | 1 | 5 | 208700 | 50 | 1 | 1 | 1 | 147000 |
| 4 | 1800000 | 2 | 6 | 908833 | 35 | 1 | 3 | 1 | 0 |
| 5 | 500000 | 1 | 5 | 166933 | 48 | 2 | 3 | 1 | 154000 |
| 6 | 500000 | 2 | 5 | 101833 | 44 | 2 | 3 | 1 | 0 |
| 7 | 500000 | 1 | 5 | 127067 | 52 | 2 | 3 | 1 | 0 |
| 8 | 1000000 | 3 | 5 | 157933 | 41 | 2 | 1 | 1 | 0 |
| 9 | 1500000 | 3 | 7 | 235464 | 30 | 1 | 3 | 1 | 33000 |
| 10 | 2000000 | 2 | 6 | 207570 | 29 | 1 | 1 | 1 | 77000 |
| 11 | 1300000 | 1 | 6 | 231222 | 45 | 1 | 2 | 1 | 80000 |
| 12 | 1200000 | 1 | 8 | 178188 | 26 | 2 | 3 | 1 | 140000 |
| 13 | 1200000 | 1 | 5 | 178200 | 50 | 1 | 1 | 1 | 56000 |
| 14 | 750000 | 2 | 10 | 37308 | 41 | 1 | 3 | 1 | 33000 |
| 15 | 4500000 | 3 | 9 | 490759 | 42 | 1 | 3 | 1 | 0 |
| 16 | 2000000 | 1 | 6 | 2428386 | 53 | 1 | 3 | 1 | 0 |
| . | . | . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . | . | . |
| . | . | . | . | . | . | . | . | . | . |
| 5503 | 800000 | 2 | 7 | 327381 | 48 | 1 | 1 | 2 | 49000 |
| 5504 | 1000000 | 2 | 5 | 272780 | 35 | 1 | 1 | 2 | 56000 |
| 5505 | 400000 | 1 | 5 | 166080 | 42 | 1 | 1 | 2 | 49000 |
| 5506 | 1000000 | 1 | 4 | 440892 | 38 | 1 | 1 | 2 | 10500 |
| 5507 | 1000000 | 1 | 4 | 553013 | 35 | 1 | 1 | 2 | 50000 |

Lampiran 2 Penjelasan Mengenai *Loss Function*

$$\begin{aligned}\text{Diketahui } \text{loss function} = \rho_{\theta} &= [\theta I(u \geq 0) + (1 - \theta)I(u < 0)]|\varepsilon| \\ &= [\theta - I(u < 0)]\varepsilon\end{aligned}$$

Dengan

ε = error dari pendugaan

$I(\varepsilon)$ = fungsi indikator yang didefinisikan

$$I(\varepsilon \geq 0) = \begin{cases} 1 & , \text{jika } u \geq 0 \\ 0 & , \text{jika } u < 0 \end{cases}$$

Sehingga :

$$\rho_{\theta} = \begin{cases} \theta u & , \text{jika } u \geq 0 \\ (\theta - 1)u & , \text{jika } u < 0 \end{cases} \quad \text{dengan } |\varepsilon| = \begin{cases} u & , \text{jika } u \geq 0 \\ -u & , \text{jika } u < 0 \end{cases}$$

Bukti:

Untuk $u \geq 0$

Persamaan (1)

$$\begin{aligned}\rho_{\theta} &= [\theta I(u \geq 0) + (1 - \theta)I(u < 0)]|\varepsilon| \\ &= [\theta I(u \geq 0) + (1 - \theta)I(u < 0)]\varepsilon \\ &= [\theta \cdot 1 + (1 - \theta)I(u < 0)]\varepsilon \\ &= [\theta + I(u < 0) - \theta I(u < 0)]\varepsilon \\ &= [\theta + (1 - I(u \geq 0)) - \theta(1 - I(u \geq 0))]\varepsilon \\ &= [\theta + (1 - 1) - \theta(1 - 1)]\varepsilon \\ &= \theta\varepsilon\end{aligned}$$

Persamaan (2)

$$\begin{aligned}\rho_{\theta} &= [\theta - I(u < 0)]\varepsilon \\ &= [\theta - (1 - I(u \geq 0))]\varepsilon \\ &= [\theta - (1 - 1)]\varepsilon \\ &= \theta\varepsilon\end{aligned}$$

Jadi terbukti bahwa:

$$\rho_{\theta} = [\theta I(u \geq 0) + (1 - \theta)I(u < 0)]|\varepsilon| = [\theta - I(u < 0)]\varepsilon, \text{ untuk } u \geq 0$$

Lampiran 2 (Lanjutan)

Untuk $u < 0$

Persamaan (1)

$$\begin{aligned}\rho_\theta &= [\theta I(u \geq 0) + (1 - \theta)I(u < 0)]|\varepsilon| \\ &= [\theta I(u \geq 0) + (1 - \theta)I(u < 0)](-\varepsilon) \\ &= [\theta \cdot 0 + (1 - \theta)I(u < 0)](-\varepsilon) \\ &= [(\theta - 1)I(u < 0)]\varepsilon \\ &= [(\theta - 1)(1 - I(u \geq 0))]\varepsilon \\ &= [(\theta - 1)(1 - 0)]\varepsilon \\ &= (\theta - 1)\varepsilon\end{aligned}$$

Persamaan (2)

$$\begin{aligned}\rho_\theta &= [\theta - I(u < 0)]\varepsilon \\ &= [\theta - (1 - I(u \geq 0))]\varepsilon \\ &= [\theta - (1 - 0)]\varepsilon \\ &= (\theta - 1)\varepsilon\end{aligned}$$

Jadi terbukti bahwa:

$$\rho_\theta = [\theta I(u \geq 0) + (1 - \theta)I(u < 0)]|\varepsilon| = [\theta - I(u < 0)]\varepsilon, \text{ untuk } u < 0$$

Sehingga dapat disimpulkan bahwa:

$$\rho_\theta = [\theta I(u \geq 0) + (1 - \theta)I(u < 0)]|\varepsilon| = [\theta - I(u < 0)]\varepsilon, \text{ untuk setiap } \varepsilon.$$

Lampiran 3 Syntax Scatter Plot

```
data=read.csv("datafixbgt.csv",header=TRUE,sep=",")

#scatterplot all#
datatable=data.frame(X1,X21,X22,X3,X4,X5,X6,X71,X72,X73,X8,Y)
pairs(datatable,col="blue",main="scatterplot")

#scatterplot per variabel#
#data dalam bentuk matriks
Y=cbind(Y)
X21=as.factor(data$X21)
X22=as.factor(data$X22)
X6=as.factor(data$X8)
X71=as.factor(data$X71)
X72=as.factor(data$X72)
X73=as.factor(data$X73)
X8=as.factor(data$X8)
X=cbind(X1,X21,X22,X3,X4,X5,X6,X71,X72,X73,X8)

#analisis kuantil regresi
library(quantreg)
kuantil0.25=rq(Y~X,tau=0.25,method="br")
summary(kuantil0.25)

#PLOT QR
plot(X1,Y,cex=.25,type="n",ylab="konsumsi rokok(Y)",xlab="Pendapatan(X1)")
points(X1,Y,cex=.5,col="blue")
abline(rq(Y~X1,tau=0.5),col="blue")
abline(lm(Y~X1),lty=2,col="red")#ols line
taus=c(0.05,0.25,0.75,0.95)
```

Lampiran 3 (Lanjutan)

```
for(i in 1:length(taus)){
  abline(rq(Y~X1,tau=taus[i]),col="black")
}

plot(X3,Y,cex=.25,type="n",ylab="konsumsi rokok(Y)",xlab="JART(X3)")
points(X3,Y,cex=.5,col="blue")
abline(rq(Y~X3,tau=0.5),col="blue")
abline(lm(Y~X3),lty=2,col="red")#ols line
taus=c(0.05,0.25,0.75,0.95)
for(i in 1:length(taus)){
  abline(rq(Y~X3,tau=taus[i]),col="black")
}

plot(X4,Y,cex=.25,type="n",ylab="konsumsi rokok(Y)",xlab="Rata-rata
pengeluaran per kapita(X4)")
points(X4,Y,cex=.5,col="blue")
abline(rq(Y~X4,tau=0.5),col="blue")
abline(lm(Y~X4),lty=2,col="red")#ols line
taus=c(0.05,0.25,0.75,0.95)
for(i in 1:length(taus)){
  abline(rq(Y~X4,tau=taus[i]),col="black")
}

plot(X5,Y,cex=.25,type="n",ylab="konsumsi rokok(Y)",xlab="Umur(X5)")
points(X5,Y,cex=.5,col="blue")
abline(rq(Y~X5,tau=0.5),col="blue")
abline(lm(Y~X5),lty=2,col="red")#ols line
taus=c(0.05,0.25,0.75,0.95)
for(i in 1:length(taus)){
  abline(rq(Y~X5,tau=taus[i]),col="black")
}
```

Lampiran 4 Syntax Perbandingan Metode

```
library(quantreg)
library(VGAM)
library(censReg)

#pembangkitan data
Pow_tobit=function(N,n,p) #n=jumlah data yg dibangkitkan
{e_pow=data.frame()
e_qr=data.frame()
e_tobit=data.frame()          #p=kuantil ke p
for (i in 1:N)
{
x1=rnorm(n,0,1)
x2=rnorm(n,0,1)
x3=rnorm(n,0,1)
e=rnorm(n,0,2)
ystar=1+x1+x2+x3+x4+x5+x6+x7+e
y=pmax(ystar,0)
yc=rep(0,n)

#estimasi quantil
quan=rq(y~x1+x2+x3+x4+x5+x6+x7, tau=p)
e_qr_temp=t(as.matrix(coef(quan)))
e_qr=rbind(e_qr,e_qr_temp)

#estimasi tobit
tobit=censReg(y~x1+x2+x3+x4+x5+x6+x7)
e_tobit_temp=t(as.matrix(coef(tobit)))
k=length(e_tobit_temp)
e_tobit_temp=t(as.matrix(e_tobit_temp[-k]))
```

```
e_tobit=rbind(e_tobit, e_tobit_temp)
}  
result=cbind(e_pow, e_qr, e_tobit)  
}  
  
compare=Pow_tobit(100,100,0.5)
```

Lampiran 5 Command Stata untuk Regresi Kuantil Tersensor

ssc install cqiv

- Untuk model regresi kuantil tersensor Gabungan (Wilayah Perkotaan dan Pedesaan):
cqiv y x1 x21 x22 x3 x4 x5 x6 x71 x72 x73 x8, exogenous quantile(10 25 50 75 90)
- Untuk model regresi kuantil tersensordi wilayah Perkotaan:
cqiv y x1 x21 x22 x3 x4 x5 x6 x71 x72 x73, exogenous quantile(25 50 75 90)
- Untuk model regresi kuantil tersensordi wilayah Pedesaan:
cqiv y x1 x21 x22 x3 x4 x5 x6 x71 x72 x73, exogenous quantile(25 50 75 90)

Lampiran 6 Output Analisis Regresi Kuantil Tersensor

cqiv y x1 x21 x22 x3 x4 x5 x6 x71 x72 x73 x8, exogenous quantile(10 25 50 75 90)
(fitting base model)

Censored quantile regression (exogenous) Number of obs = 5507
Censoring point = 0
No confidence intervals

| | Y | 10 | 25 | 50 | 75 | 90 |
|-----|-------|----------|----------|----------|----------|----------|
| x1 | | | | | | |
| | _b | 0.001727 | 0.001046 | 0.005069 | 0.009969 | 0.014545 |
| | mean | . | . | . | . | . |
| | lower | . | . | . | . | . |
| | upper | . | . | . | . | . |
| x21 | | | | | | |
| | _b | -32.321 | -12.3253 | -27.5812 | -16.885 | -8.81051 |
| | mean | . | . | . | . | . |
| | lower | . | . | . | . | . |
| | upper | . | . | . | . | . |
| x22 | | | | | | |
| | _b | -12.3977 | -1.67262 | -2.02295 | -2.38496 | -1.49203 |
| | mean | . | . | . | . | . |
| | lower | . | . | . | . | . |
| | upper | . | . | . | . | . |
| x3 | | | | | | |
| | _b | 6.759747 | 2.268525 | 5.090817 | 6.099257 | 10.73335 |
| | mean | . | . | . | . | . |
| | lower | . | . | . | . | . |
| | upper | . | . | . | . | . |
| x4 | | | | | | |
| | _b | -0.00487 | -0.00091 | 0.004397 | 0.009334 | 0.017458 |
| | mean | . | . | . | . | . |
| | lower | . | . | . | . | . |
| | upper | . | . | . | . | . |
| x5 | | | | | | |
| | _b | 0.111926 | -0.05625 | -0.44769 | -0.24076 | -0.37485 |
| | mean | . | . | . | . | . |
| | lower | . | . | . | . | . |
| | upper | . | . | . | . | . |

Lampiran 6 (Lanjutan)

| | Y | 10 | 25 | 50 | 75 | 90 |
|-------|-------|----------|----------|----------|----------|----------|
| x6 | | | | | | |
| | _b | -14.0987 | -20.4773 | -15.7674 | -16.96 | 4.646709 |
| | mean | . | . | . | . | . |
| | lower | . | . | . | . | . |
| | upper | . | . | . | . | . |
| x71 | | | | | | |
| | _b | 27.20255 | 0.815127 | 1.223801 | 1.800707 | -6.90475 |
| | mean | . | . | . | . | . |
| | lower | . | . | . | . | . |
| | upper | . | . | . | . | . |
| x72 | | | | | | |
| | _b | 1.625501 | -0.3323 | -2.00897 | 0.935681 | -0.07726 |
| | mean | . | . | . | . | . |
| | lower | . | . | . | . | . |
| | upper | . | . | . | . | . |
| x73 | | | | | | |
| | _b | 28.61763 | 0.012008 | -3.85416 | -3.3531 | -9.15691 |
| | mean | . | . | . | . | . |
| | lower | . | . | . | . | . |
| | upper | . | . | . | . | . |
| x8 | | | | | | |
| | _b | 2.215999 | -1.99489 | -4.54812 | -5.64909 | -9.88722 |
| | mean | . | . | . | . | . |
| | lower | . | . | . | . | . |
| | upper | . | . | . | . | . |
| _cons | | | | | | |
| | _b | -80.02 | -4.50685 | 17.61317 | 23.50709 | 33.0915 |
| | mean | . | . | . | . | . |
| | lower | . | . | . | . | . |
| | upper | . | . | . | . | . |

CQR Robustness Test Results

| | complete | c | pctj0 | ensorpt | pctabov~t | deltan |
|----|----------|----------|----------|---------|-----------|----------|
| 10 | 1 | 0.003348 | 0.926094 | 0 | 2.070093 | 1.133918 |
| 25 | 1 | 0.007625 | 24.714 | 0 | 77.75558 | 0.391022 |
| 50 | 1 | 0.105492 | 86.14491 | 0 | 95.53296 | 6.03594 |
| 75 | 1 | 0.331903 | 89.97639 | 0 | 100 | 26.79972 |
| 90 | 1 | 0.481771 | 90.01271 | 0 | 100 | 43.24397 |

Lampiran 6 (Lanjutan)

| | pctj1 | pctj0inj1 | inj1notj0 | obj1v | obj2v | thirdbe~r |
|----|----------|-----------|-----------|----------|----------|-----------|
| 10 | 2.015616 | 76.47059 | 72 | 20451.21 | 20398.97 | 1 |
| 25 | 75.41311 | 99.26525 | 2802 | 50969.78 | 50523.78 | 1 |
| 50 | 92.68204 | 99.78921 | 370 | 83631.16 | 83605.83 | 1 |
| 75 | 97.00381 | 99.778 | 398 | 79664.11 | 79727 | 0 |
| 90 | 97.00381 | 99.17289 | 426 | 52787.88 | 52754.71 | 1 |

. cqiv y x1 x21 x22 x3 x4 x5 x6 x71 x72 x73, exogenous quantile(25 50 75 90)
(fitting base model)

Censored quantile regression (exogenous) Number of obs = 2006
Censoring point = 0
No confidence intervals

| | y | 25 | 50 | 75 |
|-----|-------|----------|----------|----------|
| x1 | | | | |
| | _b | 0.004102 | 0.005283 | 0.010696 |
| | mean | . | . | . |
| | lower | . | . | . |
| | upper | . | . | . |
| x21 | | | | |
| | _b | -21.9064 | -14.3568 | -21.2671 |
| | mean | . | . | . |
| | lower | . | . | . |
| | upper | . | . | . |
| x22 | | | | |
| | _b | -3.96423 | -4.2424 | -7.60465 |
| | mean | . | . | . |
| | lower | . | . | . |
| | upper | . | . | . |
| x3 | | | | |
| | _b | 4.407744 | 3.440511 | 5.096841 |
| | mean | . | . | . |
| | lower | . | . | . |
| | upper | . | . | . |
| x4 | | | | |
| | _b | 0.822314 | 0.594481 | 0.460542 |
| | mean | . | . | . |
| | lower | . | . | . |
| | upper | . | . | . |

Lampiran 6 (Lanjutan)

| | | | | |
|-------|-------|----------|----------|----------|
| x5 | | | | |
| | _b | -0.12433 | -0.04778 | 0.050757 |
| | mean | . | . | . |
| | lower | . | . | . |
| | upper | . | . | . |
| x6 | | | | |
| | _b | -30.2209 | -22.0231 | -23.9408 |
| | mean | . | . | . |
| | lower | . | . | . |
| | upper | . | . | . |
| x71 | | | | |
| | _b | -1.27972 | 2.578511 | 2.437603 |
| | mean | . | . | . |
| | lower | . | . | . |
| | upper | . | . | . |
| x72 | | | | |
| | _b | -0.02767 | -1.17843 | 8.444445 |
| | mean | . | . | . |
| | lower | . | . | . |
| | upper | . | . | . |
| x73 | | | | |
| | _b | -2.01957 | -1.96525 | 2.50243 |
| | mean | . | . | . |
| | lower | . | . | . |
| | upper | . | . | . |
| _cons | | | | |
| | _b | -36.7981 | -3.34032 | 6.779411 |
| | mean | . | . | . |
| | lower | . | . | . |
| | upper | . | . | . |

CQR Robustness Test Results

| | complete | c | pctj0 | censorpt | pctabov~t | deltan |
|----|----------|----------|----------|----------|-----------|----------|
| 25 | 1 | 0.00275 | 3.838485 | 0 | 99.8006 | 9.721073 |
| 50 | 1 | 0.009967 | 38.335 | 0 | 54.18744 | 0.540751 |
| 75 | 1 | 0.103823 | 85.74277 | 0 | 100 | 10.75829 |
| 90 | 1 | 0.331481 | 89.73081 | 0 | 100 | 37.04072 |

Lampiran 6 (Lanjutan)

| | pctj1 | pctj0inj1 | inj1notj0 | obj1v | obj2v | thirdbe~r |
|----|----------|-----------|-----------|----------|----------|-----------|
| 25 | 96.75972 | 84.41558 | 1876 | 36569.87 | 8747.14 | 1 |
| 50 | 52.59222 | 96.48895 | 313 | 20985.39 | 20891.79 | 1 |
| 75 | 97.00897 | 99.47674 | 235 | 35047.76 | 34646 | 1 |
| 90 | 97.00897 | 98.22222 | 178 | 33852.38 | 33709.51 | 1 |

. cqiv y x1 x21 x22 x3 x4 x5 x6 x71 x72 x73, exogenous quantile(25 50 75 90)
(fitting base model)

Censored quantile regression (exogenous) Number of obs = 3501
Censoring point = 0
No confidence intervals

| | y | 25 | 50 | 75 | 90 |
|-----|-------|----------|----------|----------|----------|
| x1 | | | | | |
| | _b | 0.000629 | 0.007531 | 0.013151 | 0.02239 |
| | mean | . | . | . | . |
| | lower | . | . | . | . |
| | upper | . | . | . | . |
| x21 | | | | | |
| | _b | -1.53872 | -12.5866 | -9.24008 | 8.749051 |
| | mean | . | . | . | . |
| | lower | . | . | . | . |
| | upper | . | . | . | . |
| x22 | | | | | |
| | _b | -0.51877 | 0.176236 | 1.290187 | 0.189987 |
| | mean | . | . | . | . |
| | lower | . | . | . | . |
| | upper | . | . | . | . |
| x3 | | | | | |
| | _b | 1.74633 | 4.995098 | 5.871882 | 9.608563 |
| | mean | . | . | . | . |
| | lower | . | . | . | . |
| | upper | . | . | . | . |
| x4 | | | | | |
| | _b | 0.169826 | 0.460649 | 0.368602 | 0.530289 |
| | mean | . | . | . | . |
| | lower | . | . | . | . |
| | upper | . | . | . | . |

Lampiran 6 (Lanjutan)

| | Y | 25 | 50 | 75 | 90 |
|-------|-------|----------|----------|----------|----------|
| x5 | | | | | |
| | _b | -0.03569 | -0.3416 | -0.21859 | -0.36307 |
| | mean | . | . | . | . |
| | lower | . | . | . | . |
| | upper | . | . | . | . |
| x6 | | | | | |
| | _b | -0.54817 | -14.9762 | -7.03605 | -0.52907 |
| | mean | . | . | . | . |
| | lower | . | . | . | . |
| | upper | . | . | . | . |
| x71 | | | | | |
| | _b | 1.396934 | 4.181415 | 5.377321 | 10.86979 |
| | mean | . | . | . | . |
| | lower | . | . | . | . |
| | upper | . | . | . | . |
| x72 | | | | | |
| | _b | 0.478016 | 1.69548 | 3.980013 | 4.880756 |
| | mean | . | . | . | . |
| | lower | . | . | . | . |
| | upper | . | . | . | . |
| x73 | | | | | |
| | _b | 0.808612 | -3.58808 | -4.58604 | -0.66545 |
| | mean | . | . | . | . |
| | lower | . | . | . | . |
| | upper | . | . | . | . |
| _cons | | | | | |
| | _b | -9.63126 | -5.91183 | 7.059342 | 11.45422 |
| | mean | . | . | . | . |
| | lower | . | . | . | . |
| | upper | . | . | . | . |

CQR Robustness Test Results

| | complete | c | pctj0 | ensorpt | pctabov~t | deltan |
|----|----------|----------|----------|---------|-----------|----------|
| 25 | 1 | 0.007936 | 24.62154 | 0 | 73.0934 | 0.40089 |
| 50 | 1 | 0.075299 | 85.48986 | 0 | 96.8866 | 4.051673 |
| 75 | 1 | 0.299302 | 89.97429 | 0 | 100 | 23.2479 |
| 90 | 1 | 0.449264 | 89.97429 | 0 | 100 | 38.79388 |

Lampiran 6 (Lanjutan)

| | pctj1 | pctj0inj1 | inj1notj0 | obj1v | obj2v | thirdbe~r |
|----|----------|-----------|-----------|----------|----------|-----------|
| 25 | 70.89403 | 99.30394 | 1626 | 29701.86 | 28874.98 | 1 |
| 50 | 94.00171 | 99.89977 | 301 | 47566.71 | 47515.46 | 1 |
| 75 | 96.97229 | 99.71429 | 254 | 45574.61 | 45552.24 | 1 |
| 90 | 97.00086 | 99.26984 | 269 | 30532.47 | 30297.19 | 1 |

(halaman ini sengaja dikosongkan)

BIOGRAFI PENULIS



Penulis lahir di Kabupaten Cirebon, Provinsi Jawa Barat pada tanggal 04 Januari 1988 sebagai anak ketiga dari delapan bersaudara dari pasangan H.Muklas dan Hj. Tureci. Penulis menempuh pendidikan formal di SD Negeri Gamel II (Tahun 1994-2000), SLTP Negeri 1 Cirebon Barat (Tahun 2000-2003), SMA Negeri 4 Cirebon (Tahun 2003-2006).

Pada Tahun 2007, Penulis melanjutkan jenjang S1 Jurusan Statistika Universitas Padjadjaran Bandung dan selesai pada Tahun 2011. Penulis pernah bekerja pada PT Bank BNI Syariah kota Cirebon (Tahun 2012-2014). Penulis melanjutkan studi ke jenjang S2 pada semester Ganjil Tahun Akademik 2015/2016 di Program Pascasarjana Statistika FMIPA ITS Surabaya.

Segala saran, kritik, dan pertanyaan mengenai tesis ini dapat disampaikan ke penulis melalui email cintianimuklas45@gmail.com.

(Halaman ini sengaja dikosongkan)